UNIVERSIDAD NACIONAL DE CAÑETE

FACULTAD DE INGENIERÍA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete.

TESIS PARA OPTAR EL TITULO PROFESIONAL DE INGENIERO DE SISTEMAS

AUTOR(ES):

JEANICE DANIELA CHUMPITAZ LUNA

orcid.org/0009-0007-8447-624X

ANEL CECILIA HINOJOSA ROJAS

orcid.org/0000-0002-3027-1405

ASESOR:

MG. EDDYE ARTURO SANCHEZ CASTILLO

orcid.org/0000-0002-9534-6156 CO-ASESOR:

DR. CARLOS ALCIDES ALMIDÓN ORTIZ

orcid.org/0000-0003-1055-9724

FIRMA DIGITAL

Firmado digitalmente por: ALMIDON ORTIZ CARLOS ALCIDES
Motivo: Soy el autor del documento
Fecha: 20/07/2024 22:00:38-0500

LÍNEA DE INVESTIGACIÓN

CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL

San Vicente de Cañete, Perú, 2024

Firma de los jurados



Firmado digitalmente por:
VICENTE RAMOS Wagner Enoc
FAU 20491363402 soft
Motivo: En señal de
conformidad
Fecha: 15/08/2024 17:40:17-0500

Dr. Wagner Enoc Vicente Ramos
Jurado Evaluador
Facultad de Ingeniería



Firmado digitalmente por:
ASTO HUAMAN Leonidas FAU
20491363402 soft
Motivo: Doy V° B°
Fecha: 15/08/2024 17:54:11-0500

Mg. Leonidas Asto Huaman
Presidente
Jurado Evaluador
Facultad de Ingeniería

Firmado digitalmente por PACHECO PUMALEQUE ALEX ABELARDO
Motivo: Soy el autor del documento
Fecha: 06.08.2024 18:32:48 -05:00

Mg. Alex Abelardo Pacheco Pumaleque
Vocal
Jurado Evaluador
Facultad de Ingeniería

Acta de sustentación

UNIVERSIDAD NACIONAL DE CAÑETE

Código: F-M01.03.04-GT-003 Fecha de Aprobación: 16-10-2023	Versión: 04

FACULTAD DE INGENIERIA

ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO DE SISTEMAS

En la ciudad de Cañete, en la sala de sustentaciones N° 02 de la Facultad de Ingeniería, siendo las 09:00 horas, del día 21 de julio del 2024, se reunió el Jurado Evaluador y el asesor, designados por DECRETO 027-2024-UI-FI/UNDC.
Presidente(a): M. Sc. Leonidas Asto Huaman
Secretario (a): Dr. Wagner Enoc Vicente Ramos
Vocal : Mg. Alex Abelardo Pacheco Pumaleque

Asesor : Mg. Eddye Arturo Sanchez Castillo

Para el acto de sustentación de la tesis titulada **Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete**; presentada por el Bachiller (es) **Chumpitaz Luna Jeanice Daniela Y Hinojosa Rojas Anel Cecilia** para optar al Título Profesional de Ingeniero de Sistemas.

Después de concluido el acto de sustentación y luego de que el sustentante dio respuesta a las preguntas respectivas, el Jurado Evaluador calificó la sustentación o defensa de la investigación según los siguientes criterios, y promediando las notas del jurado se obtuvo:

Tesista: Chumpitaz Luna Jeanice Daniela						
Criterio	**Desaprobado (menor de 14)**	**Regular (14 -15)**	**Bueno (16-17)**	**Muy bueno (18-19)**	**Excelente (20)**	**Puntaje**
Claridad de la exposición		15				15
Dominio del tema		14				14
Visión holística del trabajo		15				15
Promedio						**15**

Tesista: Hinojosa Rojas Anel Cecilia						
Criterio	**Desaprobado (menor de 14)**	**Regular (14 -15)**	**Bueno (16-17)**	**Muy bueno (18-19)**	**Excelente (20)**	**Puntaje**
Claridad de la exposición		15				15
Dominio del tema		14				14
Visión holística del trabajo		15				15
Promedio						**15**

Para obtener la nota final de todo el proceso de investigación, conforme al artículo 100° del Reglamento General de Investigación de la Universidad Nacional de Cañete aprobado mediante Resolución N° 180-2023-UNDC/CO, se procedió a calcular la nota final:

F-M01.03.04-GT-003Toda copia de este documento, sea del entorno virtual o del documento original en físico es considerada "copia no controlada"

Tesista	Nota del Informe (valor 40%)	Nota de Sustentación (valor 60%)	Nota Final
Chumpitaz Luna Jeanice Daniela	15	15	15
Hinojosa Rojas Anel Cecilia	15	15	15

Y, establece como calificación FINAL: 15 (Quince)

- Aprobado, con nota Regular (X), Bueno (...), Muy Bueno (...), Excelente (...)
- Desaprobado (...)

Por tanto, el graduado se encuentra expedito(s) (X), impedido(s) ( ) para realizar los trámites que corresponden a la obtención del Título Profesional de Ingeniero de Sistemas.

Siendo las 09:30 Horas se dio por terminado el acto de sustentación.

Dr. Wagner Enoc Vicente Ramos **Secretario**	M. Sc. Leonidas Asto Huaman **Presidente**	Mg. Alex Abelardo Pacheco Pumaleque **Vocal**	Mg. Eddye Arturo Sanchez Castillo **Asesor**

F-M01.03.04-GT-003Toda copia de este documento, sea del entorno virtual o del documento original en físico es considerada "copia no controlada"

Dedicatoria

Queremos dedicar esté presente trabajo de investigación primeramente a Dios, quien como guía estuvo presente en el transitar de nuestras vidas. A cada uno de nuestros padres que con su cariño, esfuerzo y apoyo absoluto durante todos estos años nos permitió a llegar a cumplir nuestras metas trazadas.

Agradecimientos

Queremos expresar nuestro más sincero agradecimiento a nuestros asesores Dr. Carlos Alcides Almidón Ortiz y al Dr. Eddye Arturo Sánchez Castillo, quienes con su experiencia y orientación nos acompañaron en cada etapa de este proyecto para lograr los resultados deseados.

También deseamos agradecer a la Universidad Nacional de Cañete por proporcionarnos todos los recursos necesarios para llevar a cabo nuestra investigación.

Por último, extendemos nuestro agradecimiento a todos nuestros compañeros de la escuela profesional y a nuestra familia, por su apoyo constante en diferentes momentos de nuestras vidas. En particular, queremos mencionar a nuestros padres, quienes siempre están ahí para darnos palabras de aliento y abrazos reconfortantes que renuevan nuestras energías.

Agradecemos profundamente a todos ustedes.

Índice de contenidos

Firma de los jurados ii
Acta de sustentación iii
Dedicatoria v
Agradecimientos vi
Índice de contenidos vii
Resumen x
Abstract xii
Capítulo I: Introducción 1
1.1. Planteamiento del problema 1
1.2. Formulación del problema 2
1.1.1. Problema general 2
1.1.2. Problemas específicos 2
1.3. Objetivos 2
1.3.1. Objetivo general 2
1.3.2. Objetivos específicos 2
1.4. Justificación 3
1.4.1. Justificación teórica 3
1.4.2. Justificación practica 3
1.4.3. Justificación metodológica 3
1.5. Delimitación de la investigación 4
1.5.1. Delimitación teórica o conceptual 4
1.5.2. Delimitación temporal 4
1.5.3. Delimitación espacial 4
Capítulo II: Marco Teórico 4
2.1. Antecedentes de investigación 4
2.1.1. Artículos científicos 4
2.1.2. Tesis nacionales e internacionales 6
2.2. Bases teóricas 8
2.2.1. Variable Independiente: Sistema de Predicción 8
2.2.2. Variable Dependiente: Rendimiento Académico 11
2.3. Definición de términos básicos 12
Capítulo III: Hipótesis y Variables 13
3.1. Hipótesis 13
3.1.1. Hipótesis general 13

3.1.2. Hipótesis específicas 13
3.2. Definición conceptual de las variables 13
3.3. Operacionalización de las variables 15
Capítulo IV: Metodología 16
4.1. Enfoque de la investigación 16
4.2. Tipo de investigación 16
4.3. Nivel de investigación 16
4.4. Método de investigación 16
4.5. Diseño de investigación 16
4.6. Población, muestra y muestreo 17
4.6.1. Población 17
4.6.2. Muestra 17
4.6.3. Muestreo 17
4.7. Técnicas e instrumentos de recolección de datos 18
4.8. Técnicas estadísticas de análisis de datos 18
4.9. Consideraciones éticas 18
Capítulo V: Resultados 18
5.1. Resultados descriptivos 19
5.2. Prueba de Hipótesis 24
Capítulo VI: Discusión 35
Capítulo VII: Conclusiones 36
Capítulo VIII: Recomendaciones 37
Referencias 38
Anexos 43
Anexo 1: Matriz de consistencia 44
Anexo 2: Instrumento de recolección de datos 46
Anexo 3: Validación de Instrumentos de Investigación 48
Anexo 4: Constancia de grados y títulos de los validadores (SUNEDU) 51
Anexo 5: Bases de datos 54
Anexo 6: Autorización de la Institución para realizar la investigación 75
Anexo 7: Metodología de desarrollo de la solución 76
Anexo 8: Artículo Científico 83
Anexo 9: Constancia de similitud 91
Anexo 10: Autorización de inclusión en el repositorio de la biblioteca 92
Anexo 11: Autorización de repositorio institucional de la UNDC 93

Índice de Tablas

Tabla 1: Operacionalización de variables 15
Tabla 2: Cuadro comparativo de metodologías para minería de datos 78

Índice de Figura

Figura 1: Histograma de variación de nota 19
Figura 2:Histograma de variación de edad 19
Figura 3: Histograma de variación de Tardanza Injustificada anterior 20
Figura 4:Histograma de variación de Tardanza Justificada anterior 20
Figura 5:Histograma de variación de Inasistencia Injustificada anterior 20
Figura 6: Histograma de variación de Inasistencia Justificada anterior 21
Figura 7:Histograma de variación Nota anterior 21
Figura 8: Histograma de variación Presente..... 21
Figura 9: Box Plots 22
Figura 10: Model Comparison - Mean Squared Error 23
Figura 11: Model Comparison - R2 Score 24
Figura 12: Procesamiento de datos 24
Figura 13: Prueba de algoritmo óptimo 25
Figura 14: Prueba de algoritmo óptimo 25
Figura 15:Algoritmo de predicción: entrenamiento y creación 26
Figura 16: Resultados: aprendizaje personalizado 26
Figura 17: Valores nulos 27
Figura 18: Columnas con valores nulos 27
Figura 19: Cantidad de valores nulos. 27
Figura 20: Mapa de calor de correlación 28
Figura 21: Imputación de valores nulos 28
Figura 22: Conversión de variables categóricas 29
Figura 23: Linear Regression 30
Figura 24: Ridge Regression 30
Figura 25: Lasso Regression 31
Figura 26: Decision Tree 31
Figura 27: Random Forest 32
Figura 28: Gradient Boosting 32
Figura 29: Support Vector Machine 33
Figura 30: K – Neighbors Regressor 33
Figura 31: Neural Network 34
Figura 32: Resultado de entrenamiento 34
Figura 33: Validación cruzada 35
Figura 34: Creación de algoritmo 35
Figura 35: Insignia de la Institución Educativa 79
Figura 36: Registro de notas en Excel 81
Figura 37: Registro de asistencia en Excel 82
Figura 38: Lista de algoritmos a comparar en Colab 83

Resumen

La presente investigación tiene como objetivo predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete. Esta necesidad surgió a raíz del desempeño académico de los estudiantes, utilizando información de datos

históricos existentes en la Institución Educativa Mis Primeros Pasos. Para llevar a cabo este proceso, se empleó la metodología CRISP-DM, que consta de seis etapas: (1) Comprensión del negocio, (2) Comprensión de los datos, (3) Preparación de los datos, (4) Modelado, (5) Evaluación, y (6) Despliegue. El enfoque de la investigación fue cuantitativo, de tipo aplicado, con un diseño no experimental de tipo correlacional-causal, utilizando un tamaño de muestra de 2210 registros de notas obtenidos mediante muestreo no probabilístico por conveniencia; se utilizó el análisis documental como técnica e instrumento para registrar los datos históricos. Como resultado, se determinó que el algoritmo Gradient Boosting fue óptimo para realizar predicciones de notas, evaluado a través del error cuadrático medio, mientras que el algoritmo SVM mostró una mayor fiabilidad en las predicciones del aprendizaje de los estudiantes. En conclusión, las técnicas de minería de datos permitieron prever significativamente el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete. Se recomienda implementar un sistema de monitoreo continúo utilizando técnicas avanzadas de minería de datos para seguir evaluando y mejorando las predicciones del rendimiento académico de los estudiantes.

Palabra clave: **Gradient Boosting, metodología CRISP-DM, SVM, error cuadrático medio**

Abstract

The objective of this research is to predict the academic performance of students at the Mis Primeros Pasos Educational Institution, Cañete. This need arose as a result of the academic performance of the students, using information from existing historical data at the Mis Primeros Pasos Educational Institution. To carry out this process, the CRISP-DM methodology was used, which consists of six stages: (1) Business compression, (2) Data compression, (3) Data preparation, (4) Modeling, (5) Evaluation, and (6) Deployment. The research approach was quantitative, applied, with a non-experimental correlational-causal design, using a sample size of 2,210 records of notes obtained through non-probabilistic convenience sampling; Document analysis was used as a technique and instrument to record historical data. As a result, it was determined that the Gradient Boosting algorithm was optimal for making grade predictions, evaluated through the mean square error, while the SVM algorithm showed greater reliability in the predictions of student learning. In conclusion, data mining techniques allowed us to significantly predict the academic performance of the students of the Mis Primeros Pasos Educational Institution, Cañete. It is recommended to implement a continuous monitoring system using advanced data mining techniques to continue evaluating and improving predictions of students' academic performance.

Keyword: **Gradient Boosting, CRISP-DM methodology, SVM, mean square error**

Capítulo I: Introducción

1.1. Planteamiento del problema

En el mundo actual, la digitalización ha provocado un enorme aumento en la cantidad de datos disponibles, pero este crecimiento no siempre conduce a una mayor generación de conocimiento, como lo argumentan. En numerosos sectores académicos, científicos y empresariales, el procesamiento de datos utilizando métodos tradicionales frecuentemente resulta ser difícil, extremadamente laborioso y produce resultados superficiales e insatisfactorios [1]. Uno de los temas discutidos en el ámbito educativo es el rendimiento académico, que se enfoca principalmente en los logros o dificultades que enfrentan los estudiantes. Dado que los alumnos son pilares fundamentales de las instituciones educativas y de la sociedad en general, surge la necesidad de prever aquellos estudiantes que pueden tener un periodo educativo exitoso o complicado. Esto permitiría a la universidad ofrecer alternativas para mejorar y respaldar el progreso académico de los estudiantes [2].

Es por ello que el avance significativo de la analítica ha penetrado en el ámbito educativo, donde ahora se maneja un extenso conjunto de datos que incluye información sobre los participantes del proceso educativo. Es en este punto donde la ingeniería puede desempeñar un rol crucial al contribuir a resolver varios aspectos académicoadministrativos, como la mejora del aprendizaje, la reducción de la deserción, la disminución del abandono y la optimización del rendimiento académico [3].

En este contexto, la aplicación de tecnologías enfocadas en la extracción de datos emerge como una solución eficiente para el análisis de información, y la imperiosa tarea de elevar el estándar educativo plantea diversas estrategias para cambiar de forma dinámica la concepción de la enseñanza [4].

A nivel global, en algunas instituciones educativas de Ecuador se llevan a cabo sesiones de evaluación al finalizar un examen parcial, trimestre o año escolar para analizar el desempeño académico de los estudiantes. Este análisis se centra principalmente en determinar cuántos estudiantes han tenido éxito y cuántos no han alcanzado los estándares requeridos en cada asignatura. Sin embargo, este enfoque se limita a un análisis descriptivo básico utilizando técnicas estadísticas convencionales. Esto implica que se recopila información sobre la cantidad de estudiantes con bajo rendimiento, pero no se investigan en profundidad las causas fundamentales que provocan estos resultados negativos [5].

A nivel nacional, en Perú se ha observado que el 33.1% de los estudiantes obtienen un promedio de nota igual o menor a 10.5 en modalidad presencial, mientras que este porcentaje aumenta significativamente al 55.7% bajo la modalidad virtual.

Estas estadísticas son motivo de preocupación, ya que indican que casi la mitad de los alumnos matriculados muestran un desempeño académico deficiente. [6].

Tal es la situación de la Institución Educativa Mis Primeros Pasos, ubicada en el distrito de Mala de la Provincia de Cañete brinda servicios educativos, lo cual consta de tres niveles: inicial, primaria y secundaria.

La institución educativa, que apuesta por brindar servicios educativos, ha enfrentado algunos desafíos en materia de calidad de la educación, pues el rendimiento académico de los estudiantes el año pasado fue insuficiente en comparación con el momento de la pandemia de coronavirus de 2019.

Dicha Institución carece de una información detallada, pero sí dispersa manualmente ya sea: archivos digitales, seguimiento escolar en los docentes de la Institución. Este problema en los estudiantes podría afectar en su conducta o entre otros factores y sobre todo la pérdida del año escolar.

Sin embargo, el propósito de este estudio es predecir el desempeño académico de los estudiantes utilizando información de datos históricos existentes de los estudiantes en la institución educativa Mi primer paso.

1.2. Formulación del problema

1.1.1. Problema general

¿Cómo predice un sistema el rendimiento académico de los estudiantes de la institución educativa Mis Primeros Pasos, Cañete?

1.1.2. Problemas específicos

- ¿Qué indicadores de impacto pueden ayudar a determinar los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete?
- ¿Cuál es el algoritmo óptimo para predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete?
- ¿Como la predicción de rendimiento académico ayuda a detectar a los estudiantes que necesitan aprendizaje personalizado en la Institución Educativa Mis Primeros Pasos, Cañete?

1.3. Objetivos

1.3.1. Objetivo general

Determinar la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

1.3.2. Objetivos específicos

- Determinar los indicadores de impacto que ayuda a reconocer los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete.

- Determinar el algoritmo óptimo que genere un modelo para predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.
- Determinar una predicción al rendimiento académico que ayudará con el aprendizaje personalizado a los estudiantes de la Institución Educativa Mis Primeros Pasos.

1.4. Justificación

Según Hernández, Fernández y Mendoza [7], es crucial fundamentar cualquier estudio propuesto explicando las razones y motivos por los cuales es importante o necesario llevarlo a cabo. Además, es fundamental detallar los beneficios que se esperan obtener del estudio.

1.4.1. Justificación teórica

La justificación teórica es un tema que se aborda el cual carece de teorías o bases científicas, donde el investigador enriquece el conocimiento según al estudio de investigación que se realiza [8]. Por lo tanto, esta investigación tiene aporte tecnológico sobre el pronóstico de rendimiento académico de las Instituciones Educativas, dado que su aplicación permite tener un resultado detallado referente a aquellos factores que pueden estar siendo partícipes en la baja de notas de cada uno de los estudiantes que hay dentro de la Institución.

1.4.2. Justificación practica

La justificación práctica es aquella donde el investigador tiene la capacidad de observar un problema existente donde implementa o interviene realizando una mejora [8]. Es decir, la presente investigación considera que el rendimiento académico tiene un rol muy importante en la Institución Educativa ya que es un factor principal que da a conocer si el estudiante está aprendiendo según sus calificaciones y en qué nivel se encuentra académicamente. Es por ello que el presente proyecto dará a conocer cuáles son los factores y patrones que intervienen en el aprendizaje de los estudiantes que tienen un rendimiento académico bajo para que las autoridades de la Institución tengan conocimiento de ello y dar una solución a los casos que se pueden encontrar

1.4.3. Justificación metodológica

La justificación metodológica es aquella que permite que el investigador cree una nueva metodología para solucionar un problema existente que se encuentra en el camino, es decir, crea un nuevo instrumento o interviene de forma innovadora [8]. Los instrumentos de investigación científica son primordiales dentro de un proyecto, dicha

investigación se utilizó la recolección de datos donde fueron utilizados para detectar los factores que influyeron en el rendimiento académico de cada estudiante.

1.5. Delimitación de la investigación

1.5.1. Delimitación teórica o conceptual

El proyecto de investigación abordó un sistema predictivo como tecnología para dar a conocer cuáles son los factores que influyen en el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos y así poder dar un diagnóstico con soluciones para que puedan focalizar a aquellos estudiantes que tienen un rendimiento académico bajo. Además, se optará por utilizar la metodología CRISP-DM.

1.5.2. Delimitación temporal

El presente estudio investigativo tiene como periodo de duración no mayor a 12 semanas comprendidos entre los meses de mayo y julio del año 2024. Donde los primeros meses fueron destinados a la recolección de datos históricos los estudiantes pertenecientes al nivel secundario de la Institución Educativa Mis Primeros Pasos.

1.5.3. Delimitación espacial

El espacio que abarca la investigación será dentro de las instalaciones de la Institución Educativa Mis Primeros Pasos, el cual está ubicado en el distrito de Mala perteneciente a la provincia de Cañete, región Lima.

Capítulo II: Marco Teórico

2.1. Antecedentes de investigación

2.1.1. Artículos científicos

Según Salto-Mero y Cruz-Felipe en el año 2024, en su artículo científico titulado "Exploración del Rendimiento Académico de Estudiantes de las Carreras Economía y Turismo con Power BI en los Periodos (2021)", el objetivo fue realizar un examen del rendimiento académico de estudiantes de las carreras de Economía y Turismo en la Universidad Técnica de Manabí. Se utilizó la metodología CRISP-DM. En el estudio, una vez revisada la base de datos, se definieron las variables y filtros de datos, y se aplicaron algoritmos de aprendizaje automático (como Árbol de Decisión, Bosques Aleatorios, Redes Neuronales y Máquinas de Soporte Vectorial) utilizando Python para evaluar el desempeño académico de los estudiantes. Como resultado se determinó que el algoritmo más efectivo fue el Random Forest, el cual proporcionó resultados precisos y permitió la creación de un panel con estadísticas detalladas de los estudiantes. En conclusión, los algoritmos de aprendizaje automático demostraron ser eficientes para analizar grandes

conjuntos de datos e identificar los factores que afectan significativamente el rendimiento académico. Además, la visualización del panel en Power BI facilitó la observación y análisis claros de los datos, lo que permitió a educadores y administradores tomar decisiones bien fundamentadas para mejorar el rendimiento académico y maximizar el éxito de los estudiantes [9].

Según Quijije y Maldonado en el año 2023, en su artículo científico titulado "TÉCNICA DE MINERIA DE DATOS PARA PROCESOS EDUCATIVOS EN ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES BASADO EN UN MODELO PREDICTIVO", el propósito fue desarrollar un modelo predictivo utilizando técnicas de exploración de datos aplicadas a procesos educativos de estudiantes con necesidades educativas especiales en la Unidad Educativa "Plan Internacional". En este estudio se utilizó la metodología CRISP-DM. Como resultado, se resaltó el algoritmo J48 como sólido para la construcción de modelos predictivos del desempeño académico, con potencial aplicabilidad en otros entornos educativos. En resumen, el uso de técnicas de exploración de datos en el ámbito educativo sigue evolucionando y adaptándose a los requisitos específicos. Es fundamental obtener datos que validen los procesos y la retención del conocimiento en instituciones educativas, lo que subraya la necesidad continua de mejorar los análisis y enfoques de la exploración de datos, especialmente en aplicaciones prácticas y tareas de predicción y descripción que respalden la toma de decisiones educativas. [4].

Según Quiñones y Carrasco en el año 2020, en su artículo científico titulado "Rendimiento académico empleando minería de datos", tuvo como objetivo la predicción en el rendimiento académico de los estudiantes de la Universidad Nacimiento de Jaén de la carrera de Ingeniería de Industrias alimentarias empleando la técnica de minería de datos. Optaron por una metodología CRISP-DM. Participaron en su investigación 142 alumnos de la carrera profesional. Como resultado se obtuvo una confiabilidad mayor a 83% del rendimiento académico a través del software Weka mediante tres algoritmos de clasificación. Se concluyó que a partir de las variables identificadas se pudo determinar el rendimiento académico de los jóvenes mediante tres algoritmos de minería de datos [10].

Como último artículo tenemos a Ávila y Medina en el año 2020, en su artículo científico titulado "Minería de datos para la predicción de la deserción estudiantil en la Universidad Nacional Abierta y a Distancia". Como objetivo han planteado aplicar técnicas de minería de datos a través de informaciones recopiladas para la obtención de un modelo de predicción de la deserción estudiantil. Optaron por una metodología mixta

con un enfoque de investigación proyectiva, con una metodología de minería de datos CRISP-DM. Participaron los universitarios específicamente de la primera matrícula del 2017. Teniendo como muestra a los estudiantes matriculados en el curso de Herramientas Digitales. Se concluyó que hacer predicciones sobre deserción estudiantil optimiza la toma de decisiones, puesto que mediante grandes cantidades de información se pueden describir patrones mediante la aplicación de algoritmos de árboles de decisión [11].

2.1.2. Tesis nacionales e internacionales

Con respecto a las investigaciones realizadas a nivel nacional, Paucar en el año 2023 llevó a cabo un estudio titulado "Modelo de Aprendizaje Automático para la Predicción del Desempeño Académico de los Estudiantes Ingresantes en la Facultad de Ingeniería en Industrias Alimentarias de la UNAS". El objetivo fue desarrollar un modelo de aprendizaje automático para anticipar el rendimiento académico de los estudiantes que comienzan sus estudios en dicha facultad. En este estudio se empleó la metodología CRISP-DM junto con técnicas avanzadas de análisis de datos. El análisis se realizó utilizando una muestra de 204 registros, donde se identificaron variables críticas como la forma de ingreso, la calificación de ingreso (con mayor impacto), el género del estudiante y el número de hermanos en la familia, todos ellos relevantes para predecir el desempeño académico. Mediante pruebas como Anova de Friedman, se determinó que los modelos de Random Forest e IBK lograron una precisión del 98,4%. Específicamente, el modelo de Random Forest mostró la mayor precisión para prever el éxito académico, alcanzando un 98,34%. En conclusión, el estudio destaca la complejidad inherente al rendimiento académico de los nuevos estudiantes y enfatiza que la aplicación de la metodología CRISP-DM junto con técnicas avanzadas de análisis de datos facilita la creación de modelos que pueden predecir eficazmente dicho rendimiento. Esto proporciona una herramienta potente para utilizar datos socioeconómicos y académicos relevantes en la mejora continua de los procesos educativos y en la toma de decisiones fundamentadas [12].

Según Aronés en el año 2021, llevó a cabo un estudio titulado “PREDICCIÓN DEL RENDIMIENTO ACADÉMICO BASADO EN MACHINE LEARNING, ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS, AYACUCHO 2021”. El objetivo fue diseñar un modelo de aprendizaje automático (Machine Learning) con el propósito dar predicción en el rendimiento académico de los universitarios de la Escuela Profesional de Ingeniería de Sistemas. Optó por una investigación de tipo transversal, observacional y retrospectivo con nivel descriptivo y diseño no experimental. La población se compuso de los registros de los estudiantes de dicha Escuela Profesional de los semestres 2016-II hasta 2019-I, lo cual se tomaron los mismos datos como muestra.

El resultado obtenido fue favorable aplicando el algoritmo de regresión logística que ayudó a la predicción con un 73.6% de acuerdo a la métrica de validación de la curva ROC. Se concluye que el trabajo de investigación brinda una mejor probabilidad predictiva y toma de decisiones adecuadas en el rendimiento académico de los estudiantes de la Escuela Profesional [13].

Del mismo modo, Luna en el año 2020 realizó un estudio titulado "Implementación de un Sistema de Predicción del Desempeño Académico de los Estudiantes de Ingeniería de Sistemas de la Universidad Nacional José María Arguedas utilizando Técnicas de Minería de Datos para la Adecuada Toma de Decisiones". El objetivo fue evaluar la eficacia de un sistema predictivo para el rendimiento académico de los estudiantes universitarios de la escuela profesional de Ingeniería de Sistemas de UNAJMA mediante la minería de datos. Se empleó una metodología KDD (Knowledge Discovery in Databases) y un diseño de investigación transeccional correlacional. La población estuvo compuesta por los estudiantes de Ingeniería de Sistemas de dicha universidad, utilizando un muestreo no probabilístico. Como resultado, se determinó que el algoritmo con el menor Error Absoluto Medio (0.85) entre los construidos con 1674 datos es Kstar. Esto indica que el modelo es altamente efectivo para predecir el rendimiento académico [14].

Por otro lado, a nivel internacional, Gonzabay en el año 2023 realizó un estudio titulado "Evaluación del desempeño académico de la U.E John F. Kennedy ubicada en la parroquia José Luis Tamayo del cantón Salinas mediante técnicas de minería de datos", con el objetivo principal de aplicar la minería de datos para identificar patrones en el rendimiento académico de los estudiantes y determinar así qué estrategias de aprendizaje deben mejorarse. A lo cual tuvo como resultado predictivo gracias al dashboard interactivo haciendo uso del Power BI, en el cual las partes interesadas hicieron el monitoreo respectivo de la información identificando los factores que influyen en el rendimiento académico llegando a la búsqueda de nuevas estrategias de aprendizaje [15].

Según Mantilla (2023), en su investigación titulada "Examen del desempeño académico de los alumnos de la Institución Educativa 'Aristóteles School' empleando herramientas de Inteligencia Empresarial y explotación de datos para perfeccionar el proceso de enseñanza-aprendizaje", el propósito fue analizar el desempeño académico de los estudiantes de la institución educativa "Aristóteles School" utilizando técnicas de explotación de datos e inteligencia empresarial, además de implementar un panel de control que facilitara la identificación de áreas de mejora en el modelo educativo para asegurar un aprendizaje significativo. El estudio aplicó la metodología Cross-Industry

Standard Process for Data Mining (CRISP-DM). Como resultado del análisis, se observó que el promedio general de los estudiantes aprobados se mantuvo en un intervalo de 6,5 a 7,5, manteniendo la tendencia vista en el primer periodo del estudio. Se concluyó que la utilización de herramientas de Inteligencia Empresarial junto con la técnica de explotación de datos permitió construir modelos eficaces para prever resultados esperados en los estudiantes. Esto indica que la Inteligencia Empresarial constituye un conjunto integral de herramientas y metodologías que facilitan el análisis y la toma de decisiones estratégicas en la Institución Educativa "Aristóteles School" mediante el empleo efectivo de datos [16].

2.2. Bases teóricas

2.2.1. Variable Independiente: Sistema de Predicción

A. Definición

Un sistema de predicción es una herramienta o aplicación que utiliza datos históricos y algoritmos matemáticos para hacer estimaciones sobre eventos futuros. Estos sistemas son ampliamente utilizados en diversos campos como la meteorología, las finanzas, la salud, y también en el ámbito educativo, como en el caso del rendimiento académico [15].

Los sistemas de predicción pueden emplear diferentes enfoques de aprendizaje automático, como regresión, clasificación, redes neuronales, entre otros, para desarrollar modelos que puedan prever resultados futuros con cierto grado de precisión. Estos modelos se ajustan y mejoran continuamente a medida que se disponen de más datos y se refinan los algoritmos utilizados [16].

En el contexto de un sistema de predicción, el aprendizaje automático se utiliza para entrenar modelos con datos históricos, permitiendo que estos modelos identifiquen patrones complejos y relaciones entre variables. Una vez entrenados, estos modelos pueden hacer predicciones sobre nuevos datos o eventos futuros [17].

B. Dimensiones

Técnicas predictivas:

Es avanzada debido a su análisis sofisticado, que emplea datos históricos para el aprendizaje y datos nuevos para las predicciones. Dentro de su estructura de análisis se utilizan técnicas estadísticas y algoritmos de aprendizaje, los cuales son empleados para identificar regularidades, configuraciones o tendencias, con la finalidad de ofrecer una estimación confiable sobre un acontecimiento venidero [13].

MINERIA DE DATOS

Es el proceso de encontrar patrones relevantes, tendencias o conexiones ocultas en conjuntos de datos extensos y complejos. Este procedimiento implica el uso de métodos sofisticados de análisis de datos, estadísticas y aprendizaje automático para investigar y examinar grandes cantidades de datos con el propósito de obtener información valiosa [19].

La minería de datos en el sector de Educación implica emplear métodos y herramientas de análisis de datos para obtener conocimientos útiles de conjuntos de datos educativos. Este enfoque capacita a educadores, investigadores y responsables de políticas educativas para mejorar su comprensión del desempeño estudiantil, descubrir regularidades en el comportamiento y perfeccionar las estrategias de enseñanza [20].

Técnicas de la minería de datos:

a. Técnicas predictivas (Aprendizaje Supervisado):

   Es aquel que realiza su predicción utilizando datos históricos o anotados para obtener un resultado acumulado; es decir, se manejan datos de entrada que generarán una salida anticipada [34].

   Tipos de algoritmos:

   - Regresión Lineal: Se utiliza para estimar valores numéricos continuos utilizando variables explicativas.
   - Regresión Logística: Empleada en problemas de clasificación binaria para prever la probabilidad de pertenencia a una clase específica.
   - Máquinas de Soporte Vectorial (SVM): Aplicadas en problemas de clasificación y regresión con el fin de maximizar el margen de separación entre clases.
   - Árboles de Decisión y Bosques Aleatorios: Métodos que dividen los datos en conjuntos basados en características clave, usados en problemas de clasificación y regresión.
   - Redes Neuronales Artificiales: Modelos inspirados en el cerebro humano, utilizados en una variedad de problemas complejos como el reconocimiento de imágenes y el procesamiento del lenguaje natural.

   Estos algoritmos supervisados son cruciales en aplicaciones donde se dispone de datos etiquetados y se busca predecir con precisión

resultados futuros basados en el aprendizaje de patrones presentes en los datos de entrenamiento [33].

b. Descriptivas (Aprendizaje No supervisado):

Se basa en la premisa de que el propio sistema identifica y reconoce de manera autónoma los patrones inherentes en los datos (correlaciones estadísticas), sin depender de que los programadores etiqueten los datos de entrenamiento o especifiquen con precisión los resultados esperados. El agrupamiento es una técnica popular en la que se forman conjuntos de clusters optimizando un criterio específico [33].

Tipos de algoritmos:

- Agrupamiento (Clustering):
    - K-Means: Agrupa los datos en k conjuntos (clusters) basados en características similares.
    - Hierarchical Clustering: Organiza los datos en una jerarquía de clusters, donde los clusters pueden contener subclusters.
- Análisis de Componentes Principales (PCA): Reduce la dimensionalidad de los datos manteniendo la mayor cantidad posible de la variabilidad presente en los datos originales.
- Reducción de Dimensionalidad (t-SNE): Visualiza datos de alta dimensionalidad reduciéndolos a dos o tres dimensiones para observar patrones o estructuras subyacentes.
- Reglas de Asociación: Encuentra relaciones entre variables en grandes conjuntos de datos transaccionales para descubrir patrones como "si A entonces B".

c. Aprendizaje por refuerzo:

El programa debe ejecutar una tarea específica en un entorno que varía constantemente. Aprende mediante experimentación y equivocaciones. Explora diversas alternativas y recibe incentivos si sus acciones progresan hacia el objetivo, siendo sancionado si no lo hacen. Este método se utiliza para instruir a los ordenadores en la jugabilidad de videojuegos [32].

APRENDIZAJE AUTOMÁTICO

También denominado machine learning en inglés, es una disciplina dentro del campo de la inteligencia artificial que se dedica a crear algoritmos y metodologías que posibiliten que los sistemas informáticos aprendan y mejoren

de manera automática a partir de la experiencia, prescindiendo de la necesidad de ser programados detalladamente para cada tarea específica. Este proceso se alcanza mediante el análisis de datos y la identificación de patrones y relaciones estadísticas que pueden ser empleados para realizar predicciones o tomar decisiones [33].

Gradient Boosting es una técnica de aprendizaje automático que forma parte del conjunto más extenso de técnicas de ensemble learning, particularmente dentro de la categoría de boosting. En líneas generales, Gradient Boosting se emplea para elevar la exactitud de los modelos predictivos al unir varios modelos más básicos y menos potentes en un modelo más resistente y certero. Es una técnica poderosa y versátil en el campo del aprendizaje automático, diseñada para mejorar la precisión predictiva mediante la combinación de modelos débiles en un modelo más robusto y eficaz [34].

2.2.2. Variable Dependiente: Rendimiento Académico

A. Definición

El rendimiento académico se define como el producto de lo aprendido en una materia, en la cual se refleja mediante una calificación lo cual es influenciada por varios factores. Estos factores tanto cognitivos, emocionales y sociales están relacionados en el proceso de aprendizaje. Puesto que los estudiantes aprenden en base de las experiencias y confianza en los docentes por lo cual son el reflejo de sus calificaciones [21].

B. Dimensiones

En el presente proyecto de investigación para la variable dependiente se consideraron como dimensiones:

- Nivel Económico: En este estudio, se exploró la dimensión del nivel económico como medida económica y social de una familia. Se utilizaron como indicadores el tipo de colegio y la escala de pensión, considerando que el estudiante no recibe ningún tipo de apoyo externo [21].
- Nivel demográfico: En la presente tesis, se examinó la dimensión de nivel demográfico ya que es necesario obtener informaciones de los atributos del estudiante como la edad, sexo, lugar de residencia, entre otros [14].
- Nivel Académico: En este trabajo de investigación, se evaluó la dimensión relacionada con las habilidades del estudiante que se han desarrollado y mejorado durante el proceso de enseñanza-aprendizaje.

Este proceso permite al estudiante alcanzar un nivel de rendimiento y logros académicos durante un período o semestre determinado [16].

C. Indicadores

- Tipo de colegio:
  El tipo de colegio en donde cursó el estudiante sus estudios secundarios (estatal, privado, religioso, militar, etc.) [22]
- Edad:
  La edad del estudiante al comenzar sus estudios.
- Género:
  En relación al sexo, es uno de los factores más relevantes en la predicción del rendimiento académico, ya que se ha demostrado que existen diferencias en el estilo de aprendizaje y comportamiento entre estudiantes de género masculino y femenino. Las mujeres suelen ser más disciplinadas, capaces de mantener la concentración por períodos más largos, organizarse y mantener un nivel de motivación adecuado para enfrentar los desafíos de los estudios universitarios. [15]
- Calificación final:
  La calificación final obtenida, que se utiliza como el principal atributo para hacer predicciones. La relevancia de este atributo radica en su valor medible y utilizable como indicador del potencial académico alcanzado. [15].

2.3. Definición de términos básicos

Modelos predictivos: los modelos predictivos utilizan un conjunto de informaciones(datos) para un aprendizaje supervisado. Estos suelen utilizar regresiones lineales o no lineales, redes neuronales, árboles de decisión, entre otras [32].

Minería de datos: Son algoritmos que aplican grandes volúmenes de datos con el propósito de obtener resultados deseados [28].

Base de datos: es una recopilación de informaciones estructuradas que se almacena electrónicamente de forma ordenada para una mayor facilidad, evitando duplicidad de las mismas [25].

Metodología CRISP-DM: significa "Cross-Industry Standard Process for Data Mining", es una metodología ampliamente utilizada para proyectos de minería de datos y análisis predictivo [26].

Metodología KDD: significa "Knowledge Discovery in Databases" (Descubrimiento de Conocimiento en Bases de Datos), es un proceso integral para descubrir conocimiento útil a partir de grandes volúmenes de datos [28].

Error Absoluto Medio: Error Medio Absoluto: es una medida habitualmente empleada para evaluar la exactitud de un modelo de predicción. Se determina como el promedio de las diferencias absolutas entre los valores predichos por el modelo y los valores reales observados [32].

Patrones: se refieren a regularidades discernibles o estructuras que pueden identificarse en conjuntos de datos. Identificar patrones es fundamental porque permite extraer información significativa y tomar decisiones informadas [32].

modelo Random Forest: es un algoritmo de aprendizaje supervisado que se emplea tanto para clasificación como para regresión. Se fundamenta en la técnica de ensamblaje, donde se unen varios árboles de decisión durante el proceso de entrenamiento y predicción. [32].

Ensemble learning: se refiere a una técnica en el campo del aprendizaje automático donde se combinan múltiples modelos de predicción para mejorar la precisión y el rendimiento general del sistema predictivo [32].

Capítulo III: Hipótesis y Variables

3.1. Hipótesis

3.1.1. Hipótesis general

La aplicación de minería de datos ayuda a predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

3.1.2. Hipótesis específicas

- Los factores encontrados influyen significativamente en la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.
- Se puede determinar el algoritmo óptimo que genera un modelo predictivo en el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.
- La predicción de rendimiento académico detectó el aprendizaje personalizado que necesitan algunos estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

3.2. Definición conceptual de las variables

Variable. Sistema de Predicción

Un sistema de predicción en el contexto del rendimiento académico es una aplicación del aprendizaje automático que utiliza datos históricos y variables relevantes para prever cómo un estudiante o grupo de estudiantes pueden desempeñarse en el futuro. Este tipo de sistema utiliza algoritmos y modelos matemáticos para analizar patrones en los datos y generar predicciones basadas en esos patrones. Los datos utilizados pueden incluir información como el historial académico previo, características demográficas, resultados de pruebas estandarizadas, asistencia, entre otros [17].

Variable. Rendimiento Académico

El rendimiento académico es aquel factor supuesto que da a conocer la capacidad que tiene un estudiante, todo ello es evaluado mediante pruebas, actividades y otros que se van presentando en el transcurso del año escolar. Por otro lado, en algunos casos, se puede percibir que en los adolescentes es un cambio al cual deben adaptarse dado que el estudio es más fuerte y puede ser más estresante para algunos.[21]

3.3. Operacionalización de las variables

Tabla 1: Operacionalización de variables

TITULO: Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete				
AUTORES: JEANICE DANIELA CHUMPITAZ LUNA // ANEL CECILIA HINOJOSA ROJAS				
Variables	**Definición**	**Dimensiones**	**Indicadores**	**Instrumento**
Sistema de predicción	Es una aplicación del aprendizaje automático que utiliza datos históricos y variables relevantes para prever cómo un estudiante o grupo de estudiantes pueden desempeñarse en el futuro.	Técnica predictiva	Minería de datos	
Rendimiento Académico	Es aquel factor supuesto que da a conocer la capacidad que tiene un estudiante, todo ello es evaluado mediante pruebas, actividades y otros que se van presentando en el transcurso del año escolar.	Nivel Económico	Tipo de colegio	Base de datos
			Escala de pensión	
		Nivel Demográfico	Edad	
			Género	
			Residencia	
		Rendimiento	Calificación final al terminar el bimestre	
			Asistencia por periodo	
			Curso	

Capítulo IV: Metodología

4.1. Enfoque de la investigación

Este trabajo se centra en una metodología de investigación Cuantitativa, ya que según [7], se explica que "el enfoque cuantitativo utiliza la recolección de datos para verificar hipótesis mediante mediciones numéricas y análisis estadístico, con el objetivo de establecer patrones de comportamiento y validar teorías previamente establecidas".

4.2. Tipo de investigación

El estudio aplicado [7] es de tipo de investigación científico, está basado en la identificación y análisis de los procesos prácticos, formulación de la solución y el diseño del plan piloto para brindar soluciones en diversos campos y sectores para la toma de decisiones y una mejor eficacia brindando soluciones efectivas.

Es aplicada en base a que se determinará la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

4.3. Nivel de investigación

El nivel explicativo [8] permite identificar factores de causa y efecto de las relaciones entre variables, para el desarrollo de teorías y modelos que explican el fenómeno estudiado. En este caso, el nivel de investigación es explicativo, puesto que la necesidad se basará en identificar los cambios del sistema predictivo que influyen en la toma de decisiones del rendimiento académico de los estudiantes

4.4. Método de investigación

Teórico-analítico-sintético: se busca caracterizar la relación que existe entre los indicadores económicos, demográficos y rendimiento de los estudiantes del nivel secundaria de la Institución Educativa Mis Primeros Pasos.

Teórico lógico-histórico: evaluación de los estudiantes del nivel secundario de la Institución Educativa Mis Primeros Pasos mediante datos históricos influyentes en su rendimiento académico como: datos económicos, demográficos y rendimiento. Teóricos por modelación: se da una descomposición operacional ya que se busca saber la relación que existe en los indicadores económicos, demográficos y de rendimiento de los estudiantes del nivel secundario de la Institución Educativa Mis Primeros Pasos

4.5. Diseño de investigación

El diseño No-Experimental (transeccional) del tipo correlacional causal [7] no implica la manipulación intencional de variables independientes por parte del investigador. En su lugar, se observan y registran las variables según se presentan naturalmente en los participantes o en las situaciones bajo estudio. Los estudios

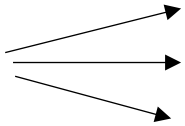
transversales recopilan datos en un solo punto en el tiempo, proporcionando una imagen instantánea de las relaciones entre variables en ese momento.

Donde se busca la relación que existe en el rendimiento académico con los indicadores económicos, demográficos y de rendimiento de cada uno de los estudiantes de la Institución.

X1

Y X2

X3



Donde:

Y: Rendimiento académico

X1: Indicador económico

X2: Indicador demográfico

X3: Indicador de rendimiento

4.6. Población, muestra y muestreo

4.6.1. Población

En la Institución educativa se tiene una población de un aproximado de 300 estudiantes, considerando que tiene los niveles de inicial, primaria y secundaria; donde se optó por trabajar con registros históricos de los estudiantes del nivel secundario desde el año 2021 hasta la actualidad.

4.6.2. Muestra

Se. realizó la clasificación de los datos históricos de los estudiantes del período 2021 al 2024 para trabajar con mayor precisión con aquellos estudiantes que se encuentran matriculados actualmente, los cuales fueron en un inicio 63 estudiantes; posteriormente se detectaron estudiantes que no tenían calificaciones en los registros brindados por la Institución, teniendo al final un total de 59 estudiantes

4.6.3. Muestreo

El muestreo no probabilístico por conveniencia es una técnica de selección de muestra en la investigación que se distingue por seleccionar participantes o elementos que son accesibles y adecuados para el investigador, sin seguir un proceso de selección aleatoria [8].

En el presente estudio de investigación se eligió el muestreo no probabilístico por conveniencia dado que la población a investigar es menor a 100. Por ello, se

seleccionaron luego de la depuración de datos 59 estudiantes con registros históricos (nivel económico, demográfico y rendimiento) para la predicción, mediante indicadores, del rendimiento académico.

4.7. Técnicas e instrumentos de recolección de datos

Se utilizó datos históricos relacionado con los estudiantes del nivel secundario de la Institución Educativa y su rendimiento académico. Centrándose en el área de matemática, donde se observó mayor dificultad, se tuvo en cuenta: datos económicos, demográficos y datos de rendimiento en los cursos de aritmética, razonamiento matemático, geometría, álgebra y trigonometría que se obtuvieron al final según los bimestres.

Por otro lado, se aplicó la técnica de extracción de datos históricos los cuales se clasificaron en una base de datos para ser aplicados con técnicas de minería de datos.

4.8. Técnicas estadísticas de análisis de datos

Las técnicas para el análisis de datos que se aplicarán en la presenten investigación son los siguiente:

a) Se empleará la caja de bigotes como herramienta gráfica para visualizar la distribución de ciertos datos importantes que se encuentran dentro de la investigación, también ayudará a identificar los valores atípicos existentes.
b) El histograma es una herramienta que se empelará para visualizar la distribución que existe entre las variables numéricas y sus rangos.
c) El mapa de correlación permite analizar la relación entre distintos factores en un conjunto de datos. Ayuda a visualizar las correlaciones que existen entre los factores donde facilita identificar las relaciones lineales positivas o negativas.

4.9. Consideraciones éticas

En esta investigación se realizará bajo aspectos éticos, se aceptaron los siguientes principios: veracidad al comunicar el propósito del estudio, autonomía al respetar la decisión de los colaboradores, confidencialidad al mantener la data en anonimato, equidad en el trato a los trabajadores, anti plagio al citar las fuentes según la norma IEEE, originalidad al plasmar las ideas del autor y sustento en el software de Turnitin..

Capítulo V: Resultados

En este capítulo se presentan los resultados que se llegaron a obtener en el proceso de investigación, donde se desarrolló el modelo predictivo más óptimo para la predicción siendo este el objetivo principal.

5.1. Resultados descriptivos

Medidas descriptivas de los indicadores. Nota final al terminar el trimestre, edad y asistencia.

En las imágenes se puede apreciar la cantidad existente en cada factor desde el año 2021 hasta la actualidad:

Nota: En la Figura 1 se puede observar un histograma que existe entre 250 y 300 calificaciones de 12 entre los cursos que se están evaluando.

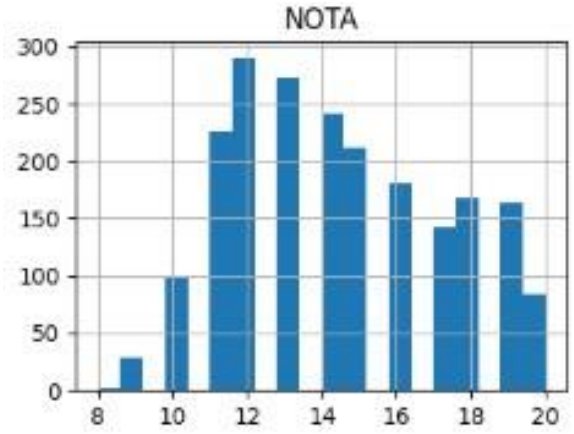


Figura 1: Histograma de variación de nota

Edad: En la Figura 2 se puede observar un histograma que se tiene un rango de edad de 12 a 15 entre los estudiantes, donde existe un aproximado de más de 800 alumnos con 12 años en los últimos años.

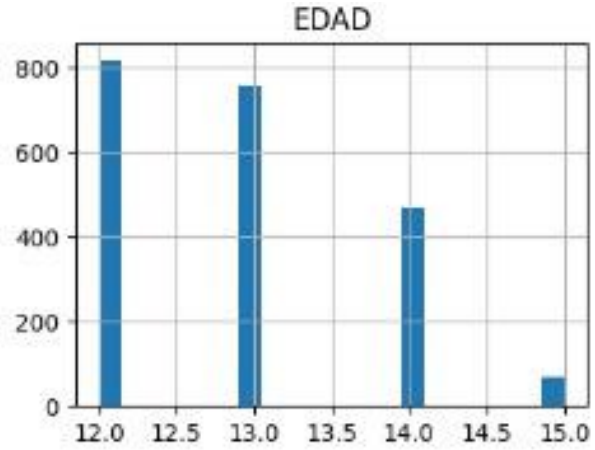


Figura 2:Histograma de variación de edad

Tardanza Injustificada anterior: En la Figura 3 se puede observar que existen más de 125 de 0 tardanzas injustificadas entre los estudiantes.

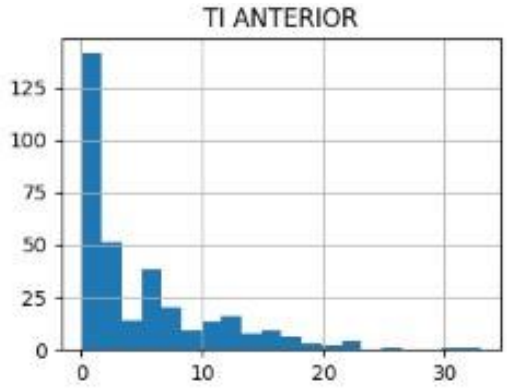


Figura 3: Histograma de variación de Tardanza Injustificada anterior

Tardanza justificada anterior: En la Figura 4 se puede observar que existe un aproximado de 300 alumnos con 0 tardanzas justificadas.

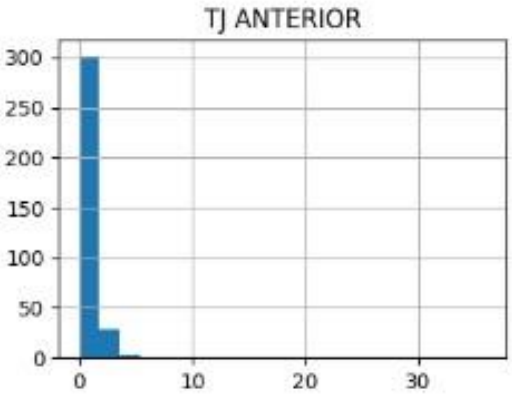


Figura 4:Histograma de variación de Tardanza Justificada anterior

Inasistencia injustificada anterior: En la Figura 5 se puede apreciar que hay más de 250 de 0 inasistencias injustificadas.

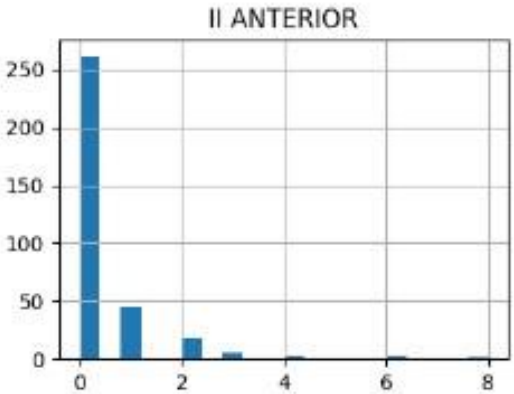


Figura 5:Histograma de variación de Inasistencia Injustificada anterior

Inasistencia justificada anterior: existen casi 250 inasistencias justificadas de 0 alumnos.

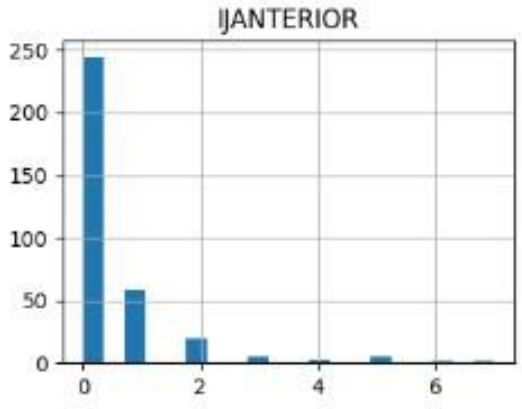


Figura 6: Histograma de variación de Inasistencia Justificada anterior

Nota anterior: se puede observar que existe entre 200 y 250 calificaciones de 12 entre los cursos que se están evaluando.

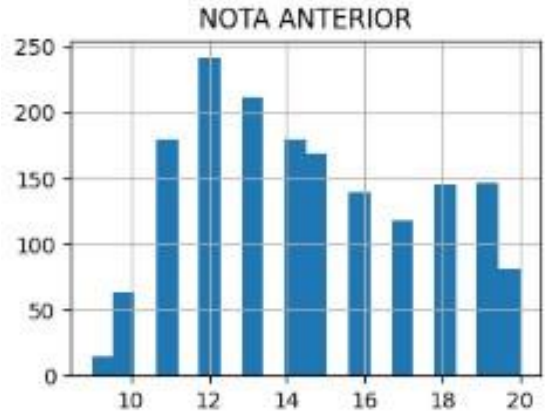


Figura 7:Histograma de variación Nota anterior

Presente: se puede observar que de 30 asistencias tiene una cantidad mayor a 30.

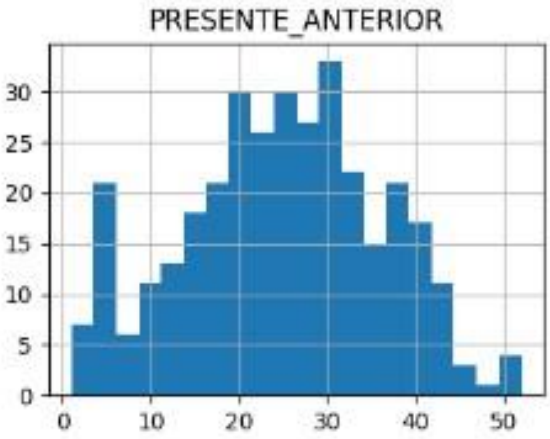


Figura 8: Histograma de variación Presente

La Figura 9 muestra el conjunto de diagramas representado en cajas donde se visualiza la distribución de los factores.

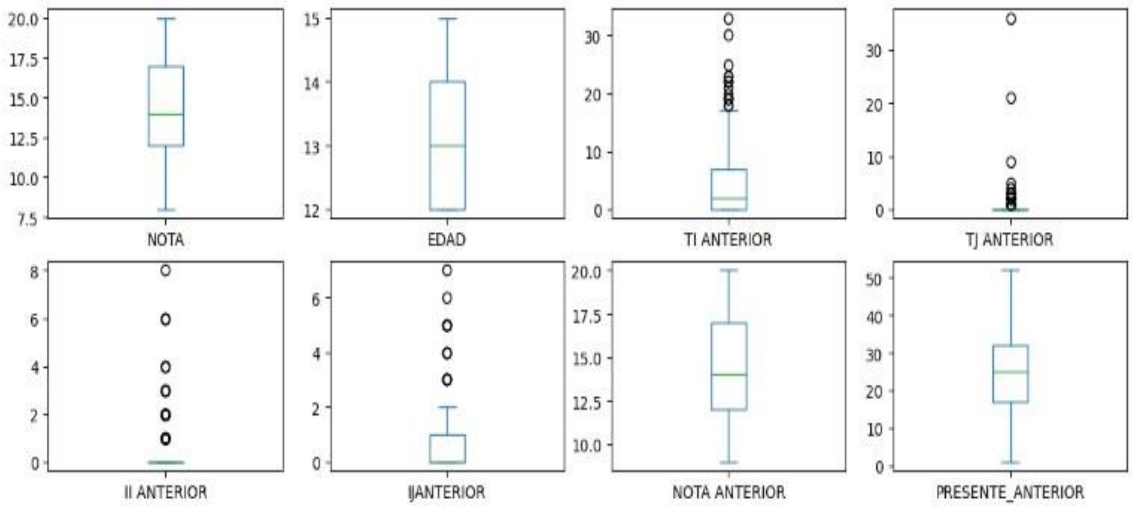


Figura 9: Box Plots

- Nota: se visualiza una mediana alrededor de 14, teniendo una distribución de notas en un rango de 12 a 17 y no existen valores atípicos.
- Edad: se tiene una mediana alrededor de 13, teniendo un rango de edades entre 12 y 14 y no existen valores atípicos.
- Tardanza injustificada anterior: se tiene una media cercana a 5, teniendo valores atípicos que están por encima de 10 lo que quiere decir que existen observaciones significativas a los demás.
- Tardanza justificada anterior: existe una mediana cerca de 2, teniendo valores atípicos donde algunos son mayores a 35.
- Inasistencia injustificada anterior: su mediana es 0, teniendo valores atípicos hasta 7.
- Inasistencia justificada anterior: su mediana es 0, teniendo valores atípicos que llegan hasta 5.
- Nota anterior: existe una mediana cerca de 14, no existen valores atípicos.
- Presente anterior: existe una variable alrededor de 30, no existen valores atípicos.

Los diagramas de caja ayudan a tener una visualización rápida de la distribución que puede existir destacando la media, los rangos y los valores atípicos existentes.

RESULTADOS DE LA PRUEBA DE MODELOS DE PREDICCIÓN

La Figura 10 muestra mediante un gráfico de barras la comparación del error cuadrático medio de los diferentes modelos, donde:

- Existen 4 modelos con un error cuadrático medio cercano a 4, siendo estos: Linear Regression, Ridge Regression, Lasso Regression y Neural Netwokrd.

- Hay 3 modelos con un error cuadrático medio cercano a 4.5, siendo estos: Random Forest, Support Vector Machine y K – Neighbors Regressor
- Un modelo con error cuadrático medio cercano a 5.5, siendo este Decision Tree.
- Un modelo con un error cuadrático medio alrededor de 3.5, siendo este Gradient Boosting.

Teniendo como modelo predictivo el Gradient Boosting, el cual tiene una mejor precisión en el rendimiento académico en comparación a los otros algoritmos.

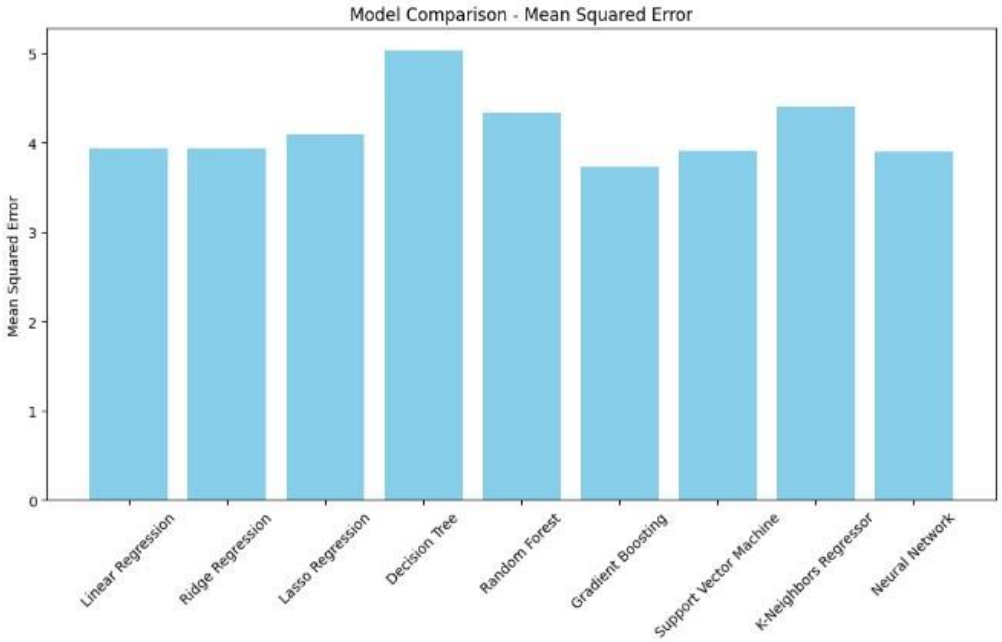


Figura 10: Model Comparison - Mean Squared Error

La Figura 11 muestra mediante un gráfico de barras la comparación del coeficiente de determinación de los diferentes modelos, donde:

- El modelo Gradient Boosting tiene un R^2 alrededor de 0.55, teniendo un mejor desempeño de comparación.
- 4 modelos tienen un R^2 alrededor de 0.5, lo cual indica que le modelo tiene una variabilidad de datos que se aproximan al 50%, los cuales son: Linear Regression, Ridge Regression, Lasso Regression y Neural Network.
- Existen dos modelos con un R^2 alrededor de 0.45 que no son tan buenos, estos son: Support Vector Machine y Random Forest.
- El modelo K – Neighbors Refressor tiene un R^2 alrededor de 0.4, siendo ligeramente menor.
- El modelo Decision Tree un R^2 alrededor de 0.3, teniendo una predicción peor que los otros modelos.

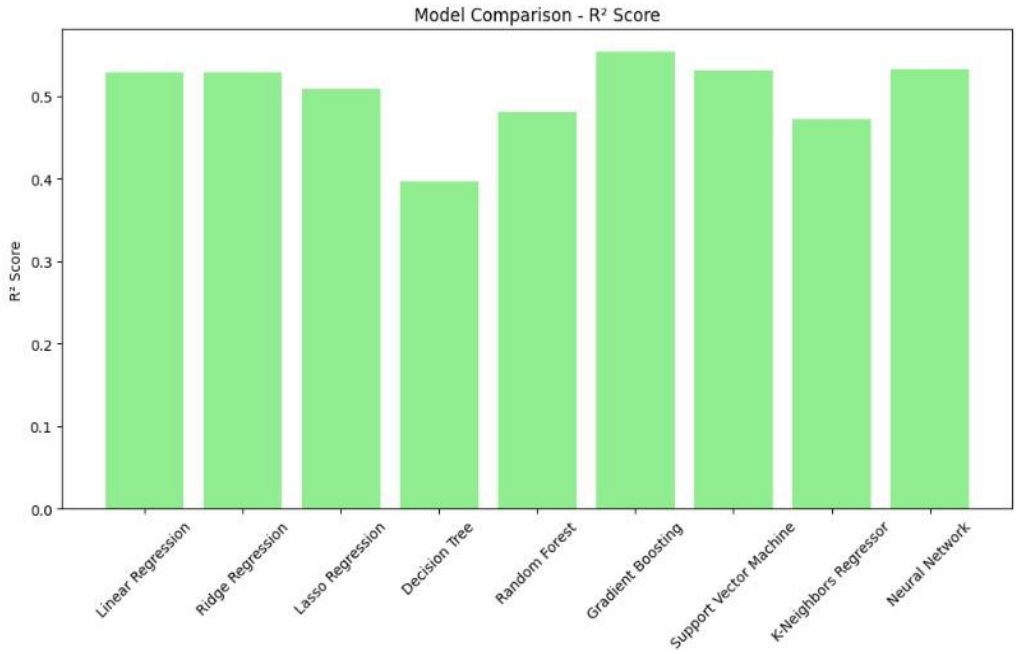


Figura 11: Model Comparison - R2 Score

El modelo Gradient Boosting tiene un R^2 Score más alto que los otros modelos, es decir, tendrá una mejor predicción del rendimiento al validar los datos. Por otro lado, el modelo de Decision Tree tiene un R^2 Score más bajo a los demás, es decir, el peor modelo para realizar una predicción.

5.2. Prueba de Hipótesis

Prueba de hipótesis con los indicadores: *Rendimiento académico y asistencia*

Para la clasificación de aprendizaje personalizado se realizó la carga de data, donde se verifica la carga de datos para realizar el procesamiento de datos y así detectar cuáles son las filas con valores nulos las que serán eliminadas posteriormente.

	EDAD	SEXO	CURSO	NOTA ANTERIOR	PRESENTE_ANTERIOR	TI ANTERIOR
0	12	1	1	NaN	NaN	NaN
1	12	1	1	NaN	NaN	NaN
2	12	1	1	NaN	NaN	NaN
3	12	0	1	NaN	NaN	NaN
4	12	0	1	NaN	NaN	NaN

	TJ ANTERIOR	II ANTERIOR	IJANTERIOR	NOTA	TIPO_APREDIZAJE
0	NaN	NaN	NaN	17	1
1	NaN	NaN	NaN	11	1
2	NaN	NaN	NaN	16	1
3	NaN	NaN	NaN	15	1
4	NaN	NaN	NaN	15	1

Figura 12: Procesamiento de datos

Prueba de hipótesis con la dimensión 3.

Para la elección del algoritmo óptimo se divide los datos en conjuntos de entrenamiento y así poder hacer la prueba de cada uno de estos; donde se inicia el entrenamiento y evaluación de cada modelo como se puede apreciar en la Figura 13 y 14:

```
Training Logistic Regression...
Logistic Regression Accuracy: 0.9447077409162717
              precision    recall  f1-score   support

           1       0.95      1.00      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.94       633
   macro avg       0.47      0.50      0.49       633
weighted avg       0.90      0.94      0.92       633

--------------------------------------
Training Decision Tree...
Decision Tree Accuracy: 0.9368088467614534
              precision    recall  f1-score   support

           1       0.95      0.99      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.94       633
   macro avg       0.47      0.49      0.48       633
weighted avg       0.89      0.94      0.92       633

--------------------------------------
Training Random Forest...
Random Forest Accuracy: 0.9447077409162717
              precision    recall  f1-score   support

           1       0.95      0.99      0.97       599
           2       0.43      0.09      0.15        34

    accuracy                           0.94       633
   macro avg       0.69      0.54      0.56       633
weighted avg       0.92      0.94      0.93       633
```

Figura 13: Prueba de algoritmo óptimo

```
--------------------------------------
Training SVM...
SVM Accuracy: 0.9462875197472354
              precision    recall  f1-score   support

           1       0.95      1.00      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.95       633
   macro avg       0.47      0.50      0.49       633
weighted avg       0.90      0.95      0.92       633

--------------------------------------
Training KNN...
KNN Accuracy: 0.9462875197472354
              precision    recall  f1-score   support

           1       0.95      1.00      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.95       633
   macro avg       0.47      0.50      0.49       633
weighted avg       0.90      0.95      0.92       633
```

Figura 14: Prueba de algoritmo óptimo

Como resultado de los algoritmos se puede observar los detalles de cada uno luego de ser evaluados, donde se percibe que la métrica que arroja es muy cercana a 1 siendo este el

algoritmo SVM con una exactitud más confiable que los otros algoritmos. Por ello, se considera un algoritmo óptimo para trabajar las predicciones.

ENTRENAR Y EVALUAR MODELO

Para la investigación se trabajó con el algoritmo SVM, el cual tuvo una mejor confiabilidad en las predicciones para así empezar a realizar la carga de data y realizar el entrenamiento del mismo. Teniendo en cuenta que la data pasó por un procedimiento, donde se tuvieron que evaluar y eliminar aquellos datos nulos para dar una mejor predicción; luego los datos se transformaron para facilitar el manejo y dividirlo en conjuntos de entrenamiento. Posteriormente se inició el clasificador del algoritmo para empezar con el entrenamiento y la prueba de este, dando como resultado la precisión, el Accuracy, entre otros, como se puede percibir en la Figura 15:

```
SVM Accuracy: 0.9462875197472354
              precision    recall  f1-score   support

           1       0.95      1.00      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.95       633
   macro avg       0.47      0.50      0.49       633
weighted avg       0.90      0.95      0.92       633
```

Figura 15:Algoritmo de predicción: entrenamiento y creación

EDAD	SEXO	CURSO	NOTA ANTERIOR	PRESENTE_ANTERIOR	TI ANTERIOR	TJ ANTERIOR	II ANTERIOR	IJANTERIOR	NOTA	TIPO_APREDIZAJE	PREDICCION_TIPO_APRENDIZAJE
12	1	1	14	24	4		0		17	1	1
12	1	1	14	24	4		0		11	1	1
12	1	1	14	24	4		0		16	1	1
12	0	1	14	24	4		0		15	1	1
12	0	1	14	24	4		0		15	1	1

Figura 16: Resultados: aprendizaje personalizado

Terminado de realizar el modelo de predicción con el algoritmo de confiabilidad se dio la predicción de aprendizaje de los estudiantes, donde se considera dos grupos en escalas: 1 es aquel alumno aprobado con calificación mayor a 10 y 2 es aquel alumno desaprobado con calificación menor a 11. Por lo tanto, aquellos alumnos que están desaprobados se consideran como alumnos que necesitan un aprendizaje personalizado para tener una mejora en su nota.

Prueba de hipótesis con los indicadores de impacto: edad, sexo, curso, calificación al finalizar el trimestres y asistencia.

Exploración de datos

	EDAD	SEXO	CURSO	NOTA ANTERIOR	PRESENTE_ANTERIOR	TI ANTERIOR
0	12	1	1	NaN	NaN	NaN
1	12	1	1	NaN	NaN	NaN
2	12	1	1	NaN	NaN	NaN
3	12	0	1	NaN	NaN	NaN
4	12	0	1	NaN	NaN	NaN

	TJ ANTERIOR	II ANTERIOR	IJANTERIOR	NOTA	TIPO_APREDIZAJE
0	NaN	NaN	NaN	17	1
1	NaN	NaN	NaN	11	2
2	NaN	NaN	NaN	16	1
3	NaN	NaN	NaN	15	1
4	NaN	NaN	NaN	15	1

Figura 17: Valores nulos

Se realiza una exploración de datos para encontrar aquellas casillas que están vacías o nulas. En la Figura 18 se visualiza los factores que tienen valores nulos después de la exploración.

0	EDAD	2110 non-null	int64
1	SEXO	2110 non-null	int64
2	CURSO	2110 non-null	int64
3	NOTA ANTERIOR	1690 non-null	float64
4	PRESENTE_ANTERIOR	337 non-null	float64
5	TI ANTERIOR	337 non-null	float64
6	TJ ANTERIOR	337 non-null	float64
7	II ANTERIOR	337 non-null	float64
8	IJANTERIOR	337 non-null	float64
9	NOTA	2110 non-null	int64
10	TIPO_APREDIZAJE	2110 non-null	int64

Figura 18: Columnas con valores nulos

Después de realizar la exploración de datos se detectan cuáles son las columnas que tienen valores nulos y que se quitarán más adelante.

None	
EDAD	0
SEXO	0
CURSO	0
NOTA ANTERIOR	420
PRESENTE_ANTERIOR	1773
TI ANTERIOR	1773
TJ ANTERIOR	1773
II ANTERIOR	1773
IJANTERIOR	1773
NOTA	0
TIPO_APREDIZAJE	0

Figura 19: Cantidad de valores nulos.

Se tiene como resultado la detección de cantidad con valores nulos en cada columna como se puede apreciar en la Figura 19.

Prueba de hipótesis con los indicadores de nota al final de trimestre y asistencia.

En el mapa de calor se pudo obtener el impacto que puede tener cada dato obtenido, dando a conocer cuáles son los que ayudarán a predecir el rendimiento académico del estudiante. Se puede observar en la Figura 20 que la Nota anterior es la que tendrá más uso para realizar una predicción real.

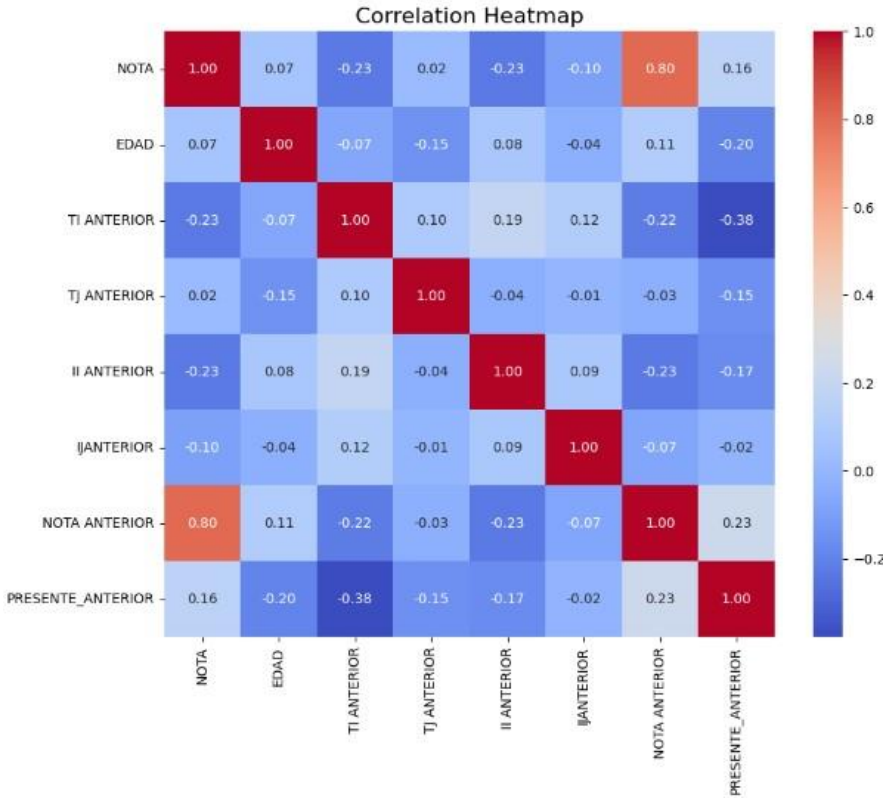


Figura 20: Mapa de calor de correlación

TRATAMIENTO DE DATOS

Se hace uso de la media para imputar los valores nulos existentes en la data para ser eliminados y permitan una mejor predicción, luego se realiza una verificación para comprobar que no existan datos nulos, así como podemos apreciar en la Figura 21:

```
EDAD                0
SEXO                0
CURSO               0
NOTA ANTERIOR       0
PRESENTE_ANTERIOR   0
TI ANTERIOR         0
TJ ANTERIOR         0
II ANTERIOR         0
IJANTERIOR          0
NOTA                0
TIPO_APREDIZAJE     0
```

Figura 21: Imputación de valores nulos

Conversión de variables

Se utiliza el one – hot encoding para convertir variables categóricas en un formato numérico binario el permita una mejor compresión de datos; como se puede visualizar en la Figura 22 las siguientes variables categóricas se convirtieron:

- Sexo: Masculino = y Femenino =
- Curso:
- Tipo de aprendizaje:

	EDAD	NOTA ANTERIOR	PRESENTE_ANTERIOR	TI ANTERIOR	TJ ANTERIOR	\
0	12.0	14.617751	24.878338	4.810089	0.554896	
1	12.0	14.617751	24.878338	4.810089	0.554896	
2	12.0	14.617751	24.878338	4.810089	0.554896	
3	12.0	14.617751	24.878338	4.810089	0.554896	
4	12.0	14.617751	24.878338	4.810089	0.554896	

	II ANTERIOR	IJANTERIOR	NOTA	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0	\
0	0.397626	0.477745	17.0	False	True	True	False	
1	0.397626	0.477745	11.0	False	True	True	False	
2	0.397626	0.477745	16.0	False	True	True	False	
3	0.397626	0.477745	15.0	True	False	True	False	
4	0.397626	0.477745	15.0	True	False	True	False	

	CURSO_3.0	CURSO_4.0	CURSO_5.0	TIPO_APREDIZAJE_1.0	TIPO_APREDIZAJE_2.0
0	False	False	False	True	False
1	False	False	False	False	True
2	False	False	False	True	False
3	False	False	False	True	False
4	False	False	False	True	False

Figura 22: Conversión de variables categóricas

PRUEBA DE MODELOS DE PREDICCIÓN

Para encontrar el modelo de predicción se realizaron pruebas con diversos algoritmos para detectar cuál de todos tenían una mejor confiabilidad al momento de dar la predicción, es decir, cuál se acercaba más a una nota real del alumno. Con las pruebas se llegó a tener diversas notas predictivas con los diferentes algoritmos, notas que variaron desde salir como alumno aprobado hasta salir desaprobado.

Se emplearon algoritmos para la prueban, teniendo como resultado, los siguientes:

- Linear Regression (Figura 23)
- Ridge Regression (Figura 24)
- Lasso Regression (Figura 25)
- Decision Tree (Figura 26)
- Random Forest (Figura 27)
- Gradient Boosting (Figura 28)
- Support Vector Machine (Figura 29)
- K – Neighbors Regressor (Figura 30)
- Neural Network (Figura 31)

Linear Regression - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	12.752778
1986	False	True	False	12.441310
420	True	False	False	14.034572
527	False	False	True	12.286648
196	False	False	False	11.353228
744	False	False	True	16.251456
1826	True	False	False	18.026407
324	False	False	False	12.327203
1449	False	False	False	13.064215
634	False	True	False	16.464846

Figura 23: Linear Regression

Ridge Regression - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	12.753783
1986	False	True	False	12.440924
420	True	False	False	14.034219
527	False	False	True	12.286877
196	False	False	False	11.353820
744	False	False	True	16.251407
1826	True	False	False	18.025778
324	False	False	False	12.329765
1449	False	False	False	13.064430
634	False	True	False	16.464209

Figura 24: Ridge Regression

Lasso Regression - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	13.325600
1986	False	True	False	12.666448
420	True	False	False	13.984752
527	False	False	True	12.666448
196	False	False	False	12.007296
744	False	False	True	15.962208
1826	True	False	False	17.280512
324	False	False	False	13.325600
1449	False	False	False	13.325600
634	False	True	False	15.962208

Figura 25: Lasso Regression

Decision Tree - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	14.000000
1986	False	True	False	12.500000
420	True	False	False	14.666667
527	False	False	True	12.500000
196	False	False	False	11.000000
744	False	False	True	16.000000
1826	True	False	False	18.571429
324	False	False	False	16.000000
1449	False	False	False	13.000000
634	False	True	False	19.666667

Figura 26: Decision Tree

Random Forest - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	13.207489
1986	False	True	False	12.708831
420	True	False	False	14.614714
527	False	False	True	12.542060
196	False	False	False	10.968410
744	False	False	True	15.833690
1826	True	False	False	18.615384
324	False	False	False	14.320333
1449	False	False	False	12.899766
634	False	True	False	19.586917

Figura 27: Random Forest

Gradient Boosting - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	13.453182
1986	False	True	False	13.358483
420	True	False	False	13.652587
527	False	False	True	12.548248
196	False	False	False	11.444076
744	False	False	True	16.632552
1826	True	False	False	18.546389
324	False	False	False	11.918431
1449	False	False	False	13.016566
634	False	True	False	17.069826

Figura 28: Gradient Boosting

Support Vector Machine - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	12.970242
1986	False	True	False	12.314318
420	True	False	False	13.918108
527	False	False	True	12.372703
196	False	False	False	11.512170
744	False	False	True	16.305419
1826	True	False	False	17.969497
324	False	False	False	13.200915
1449	False	False	False	13.157729
634	False	True	False	16.371083

Figura 29: Support Vector Machine

K-Neighbors Regressor - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	13.6
1986	False	True	False	12.2
420	True	False	False	14.4
527	False	False	True	12.0
196	False	False	False	11.2
744	False	False	True	16.6
1826	True	False	False	19.2
324	False	False	False	12.2
1449	False	False	False	12.6
634	False	True	False	16.8

Figura 30: K – Neighbors Regressor

Neural Network - Primeras predicciones:

	EDAD	TI ANTERIOR	TJ ANTERIOR	NOTA ANTERIOR	PRESENTE_ANTERIOR
544	12.0	6.000000	0.000000	13.0	26.000000
1986	12.0	4.810089	0.554896	12.0	24.878338
420	14.0	4.810089	0.554896	14.0	24.878338
527	14.0	4.810089	0.554896	12.0	24.878338
196	13.0	4.810089	0.554896	11.0	24.878338
744	12.0	4.810089	0.554896	17.0	24.878338
1826	13.0	4.810089	0.554896	19.0	24.878338
324	14.0	20.000000	0.000000	13.0	5.000000
1449	13.0	3.000000	0.000000	13.0	25.000000
634	14.0	4.810089	0.554896	17.0	24.878338

	II ANTERIOR	IJANTERIOR	SEXO_0.0	SEXO_1.0	CURSO_1.0	CURSO_2.0
544	3.000000	0.000000	True	False	True	False
1986	0.397626	0.477745	True	False	False	False
420	0.397626	0.477745	False	True	False	False
527	0.397626	0.477745	False	True	False	False
196	0.397626	0.477745	True	False	False	True
744	0.397626	0.477745	True	False	False	False
1826	0.397626	0.477745	False	True	False	False
324	1.000000	5.000000	False	True	True	False
1449	0.000000	0.000000	False	True	True	False
634	0.397626	0.477745	False	True	False	False

	CURSO_3.0	CURSO_4.0	CURSO_5.0	Predicción
544	False	False	False	12.371172
1986	False	True	False	12.398722
420	True	False	False	13.943279
527	False	False	True	12.358089
196	False	False	False	11.456932
744	False	False	True	16.057362
1826	True	False	False	18.309253
324	False	False	False	12.412433
1449	False	False	False	12.839771
634	False	True	False	16.511166

Figura 31: Neural Network

RESULTADOS DEL ENTRENAMIENTO

La Figura 32 muestra una tabla con los resultados del entrenamiento de diversos modelos de algoritmos donde se evalúan dos métricas: MSE, el cual mide el promedio de errores al cuadrado entre los valores que se llegan a predecir y los valores que son reales; y, R^2, es el encargado de medir la proporción de variación de la variable y tiene que ser más cercana a 1 que quiere decir que tiene un mejor ajuste del modelo de datos.

	Model	MSE	R²
0	Linear Regression	3.932223	0.528180
1	Ridge Regression	3.932183	0.528185
2	Lasso Regression	4.095871	0.508544
3	Decision Tree	5.025271	0.397027
4	Random Forest	4.328178	0.480670
5	Gradient Boosting	3.724235	0.553136
6	Support Vector Machine	3.905293	0.531411
7	K-Neighbors Regressor	4.403160	0.471673
8	Neural Network	3.895922	0.532535

Figura 32: Resultado de entrenamiento

El modelo Gradient Boosting presenta un rendimiento general y mejor para realizar la predicción que se quiere, teniendo un MSE de 3.724235 y un R^2 de 0.553136, lo que indica que es un modelo que se ajusta a los datos en la investigación.

	Model	CV Mean MSE
0	Linear Regression	4.215487
1	Ridge Regression	4.214888
2	Lasso Regression	4.141853
3	Decision Tree	5.394054
4	Random Forest	4.726612
5	Gradient Boosting	4.087936
6	Support Vector Machine	4.098591
7	K-Neighbors Regressor	5.002673
8	Neural Network	4.291799

Figura 33: Validación cruzada

La Figura 34 muestra los resultados de la validación cruzada en distintos modelos utilizando el error cuadrático medio promedio para la evaluación del rendimiento académico en distintos modelos con múltiples particiones de conjuntos de la data.

```
MSE en datos de prueba: 3.7244
R² en datos de prueba: 0.5531

Validación cruzada - Mean MSE: 4.0889 +/- 0.8369

RESULTADOS DE ENTRENAMIENTO
              Model       MSE        R²  CV Mean MSE
0  Gradient Boosting  3.724358  0.553121     4.088924
```

Figura 34: Creación de algoritmo

Capítulo VI: Discusión

Respecto a los indicadores de impacto

Según los resultados de este estudio, como lo muestra el mapa de calor se pudo obtener el impacto que puede tener cada dato obtenido, dando a conocer cuáles son los que ayudarán a predecir el rendimiento académico del estudiante, dicho esto, Nota anterior es la que tendrá más uso para realizar una predicción real.

Respecto a los indicadores del algoritmo óptimo

Según los hallazgos de esta investigación se determinó a través del error cuadrático medio el algoritmo Gradient Boosting tiene un R^2 alrededor de 0.55 el cual tiene una mejor precisión en el

rendimiento académico en comparación a los otros algoritmos. Mientras que el modelo Decision Tree un R^2 alrededor de 0.3, teniendo una predicción peor que los otros modelos para realizar una predicción. Como lo menciona [] el algoritmo Gradient Boosting se emplea para elevar la exactitud de los modelos predictivos al unir varios modelos más básicos y menos potentes en un modelo más resistente y certero.

Respecto a los indicadores del tipo de aprendizaje

Según los resultados del tercer indicador, queda comprobado que el algoritmo SVM (Máquinas de Soporte Vectorial) tuvo una mejor confiabilidad en las predicciones para así empezar a realizar la carga de data y realizar el entrenamiento del mismo. con una exactitud más confiable que los otros algoritmos. Por ello, se considera un algoritmo óptimo para trabajar las predicciones. Como lo menciona [] el algoritmo SVM (Máquinas de Soporte Vectorial) se utiliza principalmente para encontrar un hiperplano en un espacio multidimensional que clasifique de manera óptima los puntos de datos.

Respecto al Objetivo General

El objetivo general planteado en esta investigación fue determinar la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete, se hallaron resultados positivos que se explicaron anteriormente, podemos afirmar que se logró cumplir satisfactoriamente el objetivo principal propuesto en esta investigación.

En la primera métrica (Indicadores de impacto) se determinó que Nota Anterior tuvo una mayor influencia para la realización de la predicción, en la segunda métrica se determinó que el algoritmo Gradient Boosting fue el más óptimo para la predicción de notas y en la tercera métrica se determinó que el algoritmo SVM (Máquinas de Soporte Vectorial) fue el más óptimo para el tipo de aprendizaje.

En conclusión, se afirma que la aplicación de minería de datos ayuda a predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

Capítulo VII: Conclusiones

Se concluyó que:

1. Se determinó que las técnicas de minería de datos predicen significativamente el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

2. Se identificó mediante un mapa de calor el indicador clave para predecir el rendimiento académico, donde la nota es la que más influencia tiene.

3. Se determinó que el algoritmo con mejor precisión para predecir el rendimiento académico de los estudiantes es el modelo Gradient Boosting que presenta un rendimiento general y mejor para realizar la predicción que se quiere, teniendo un MSE de 3.724235 y un R^2 de 0.553136, lo que indica que es un modelo que se ajusta a los datos en la investigación en comparación a los otros algoritmos.

4. Se determinó que el algoritmo SVM tuvo una mejor confiabilidad en las predicciones y con mayor exactitud que los otros algoritmos, lo cual se pudo identificar a los estudiantes que necesita un aprendizaje personalizado de la materia.

Capítulo VIII: Recomendaciones

1. Se recomienda implementar un sistema de monitoreo continúo utilizando técnicas avanzadas de minería de datos para seguir evaluando y refinando las predicciones del rendimiento académico de los estudiantes.

2. Se sugiere incorporar más atributos relacionadas con el ambiente educativo del estudiante para enriquecer la información disponible y ampliar las posibilidades en futuros proyectos similares. Esto permitirá desarrollar más modelos predictivos mediante la aplicación de diversas técnicas de minería de datos basadas en las variables propuestas, lo cual asegura una mejora en la precisión de las predicciones obtenidas.

3. Se sugiere utilizar métodos de análisis de datos avanzados como agrupamiento, análisis de regresión, clasificación y descubrimiento de asociaciones para detectar patrones, tendencias y relaciones en los datos educativos. Estos análisis pueden contribuir a una comprensión más profunda de los elementos que impactan en el desempeño académico y en la efectividad de los enfoques educativos.

4. Se sugiere Capacitar a los docentes y personal educativo para que utilicen de manera efectiva las predicciones del modelo SVM en el desarrollo e implementación de programas de aprendizaje personalizado. Proporciona orientación sobre cómo interpretar

y aplicar las recomendaciones del modelo para mejorar el apoyo educativo individualizado.

Referencias

[1] D.C. Chamorro-Sangoquiza, A.M. Vargas-Muñoz, A.C. Umaquinga-Criollo, M.A. Becerra y D.H. Peluffo-Ordoñez, "Estudio comparativo de técnicas de minería de datos para develar patrones de desempeño académico en enseñanza media", *RISTI*, pp. 455-468, junio de 2020.

[2] J.D. Garcia Dionisio, "Machine Learning para predecir el rendimiento académico de los estudiantes universitarios", tesis de Licenciado, Universidad César Vallejo, Lima, UCV, Perú, 2021.

[3] L.E. Contreras, H.J. Fuentes y J.I. Rodríguez, "Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático", *SCIELO*, vol. 13(5), pp. 233-246, octubre de 2020.

[4] H.B. Quijije Quiroz y K. Maldonado Zuñiga, " TÉCNICA DE MINERÍA DE DATOS PARA PROCESOS EDUCATIVOS EN ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES BASADO EN UN MODELO PREDICTIVO", *PENTACIENCIAS*, vol. 5, pp. 205-217, septiembre de 2023.

[5] M.A. Gonzabay Panchana, "Evaluación del rendimiento académico de la U.E John F. Kennedy situada en la parroquia José Luis Tamayo del cantón Salinas aplicando técnicas de minería de datos", tesis de Maestría, Universidad Estatal Península de Santa Elena, Santa Elena, UPSE, Ecuador, 2023.

[6] L.E. Contreras, H.J. Fuentes y J.I. Rodríguez, "Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático", *SCIELO*, vol. 13(5), pp. 233-246, octubre de 2020.

[7] R. Hernández Sampieri, C. Fernández Collado y P. Baptista Lucio. *Metodología de*

la Investigación. México: McGrawHill, 2019.

[8] J. Arias-Gonzales, *DISEÑO Y METODOLOGÍA DE LA INVESTIGACIÓN*, Arequipa: ResearchGate, 2021.

[9] J.S. Salto-Mero y M.R Cruz-Felipe, "Análisis del Rendimiento Académico de Estudiantes de las Carreras Economía y Turismo con Power BI en los Periodos (2021)", *Digital Publisher*, vol. 5, pp. 762-772, enero de 2024.

[10] L. Quiñones Huatangari y Y.L. Carrasco Vega, " Rendimiento académico empleando minería de datos", *Espacios*, vol. 41(44), pp. 277-285, noviembre de 2020.

[11] R. Rojas Pari, "Modelo de Aprendizaje Automático Supervisado para Identificar Patrones de Bajo Rendimiento Académico en los Ingresantes al Instituto de Educación Superior Pedagógico Público – Juliaca ". tesis de Licenciado, Universidad Peruana Unión, Juliaca UPEU, Perú, 2021.

[12] W.G. Paucar Palomino, "MODELO DE APRENDIZAJE AUTOMATICO PARA LA PREDICCION DEL RENDIMIENTO ACADEMICO DE LOS ALUMNOS INGRESANTES EN LA FACULTAD DE INGENIERIA EN INDUSTRIAS ALIMENTARIAS DE LA UNAS", tesis de Licenciado, Universidad Nacional Agraria de la Selva, Tingo Maria, UNAS, Perú, 2023.

[13] E. Aronés Ayala, " PREDICCIÓN DEL RENDIMIENTO ACADÉMICO BASADO EN MACHINE LEARNING, ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS, AYACUCHO 2021", tesis de Licenciado, Universidad Nacional de San Cristóbal de Huamanga, Ayacucho, UNSCH, Perú, 2021.

[14] N.H. Luna Reyes, "IMPLEMENTACIÓN DE UN SISTEMA DE PREDICCIÓN DEL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES DE INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD NACIONAL JOSÉ MARÍA ARGUEDAS UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS PARA LA ADECUADA

TOMA DE DECISIONES", tesis de Licenciado, Universidad Nacional José María Arguedas, Apurimac, UNAJMA, Perú, 2020.

[15] M.A. Gonzabay Panchana, "Evaluación del rendimiento académico de la U.E John F. Kennedy situada en la parroquia José Luis Tamayo del cantón Salinas aplicando técnicas de minería de datos.", tesis de Magister, Universidad Estatal Península de Santa Elena, Santa Elena, UPSE, Ecuador, 2023.

[16] L.A. Mantilla Galarza, " Análisis del rendimiento académico de los estudiantes de la Unidad Educativa "Aristóteles School" aplicando herramientas de Business Intelligence y minería de datos para mejorar el proceso de enseñanza aprendizaje.", tesis de Magister, Universidad de las Fuerzas Armadas, Santa Elena, ESPE, Ecuador, 2023.

[17] «Qué es Modelo predictivo,» ARIMETRICS, [En línea]. Available: https://www.arimetrics.com/glosario-digital/modelo-predictivo. [Último acceso: 23 mayo 2024].

[18] K. J. Paredes-Robles, J. Vélez-Pincay y L. R. Marin-Llaver, «Estrategia Educativa para mejorar el rendimiento académico en estudiantes de bachillerato.,» 15 septiembre 2023. [En línea]. [Último acceso: mayo 2024].

[19] prom peru, SOFTWARE Minería de Datos y Almacenamiento, 2021.

[20] S. PIO ALVAREZ, «MINERÍA DE CALIDAD DE DATOS: APLICACIÓN DE TÉCNICAS DE MINERÍA DE DATOS PARA LA EVALUACIÓN DE LA CALIDAD DE LOS DATOS,» 2018. [En línea]. Available: https://www.colibri.udelar.edu.uy/jspui/bitstream/20.500.12008/25468/1/PIO18.pdf. [Último acceso: mayo 2024].

[21] A. J. VILLAO BALÓN, «Aplicación de técnicas de minería de datos para predecir el desempeño académico de los estudiantes de la escuela ‘Lic. Angélica Villón L.,»

2021. [En línea]. Available: https://repositorio.upse.edu.ec/bitstream/46000/6493/1/UPSE-TTI-2021-0044.pdf. [Último acceso: mayo 2024].

[22] A. Müller. y S. Guido (2019). Introduction to Machine Learning with Python. Boston: OREILLY.

[23] M. Pojon (2020). Using Machine Learning to predict student performance. University of Tampere

[24] M. Orallo, C. Ramírez (s. f.) "Minería de Datos y Extracción de Conocimiento de Bases de Datos"., De: http://users.dsic.upv.es/~jorallo/docent/doctorat/t2a.pdf

[25] A. Rico Páez (24 de Enero-junio de 2022). "Modelos preditivos progressivos de desempenho acadêmico de estudantes". 12(24). https://www.ride.org.mx/index.php/RIDE/article/view/1196

[26] E. Taya Acosta. "Modelo de mineria de datos para evaluar el efecto del uso del aula virtual sobre el rendimiento academico de los estudiantes de la facultad de ingenieria de la universidad nacional Jorge Basadre Grohmann de Tacna, en Tiempos de pandemia,2020". tesis de Magister, Universidad Nacional Jorge Basadre Grohmann, UNJBG, Tacna, 2021.

[27] J. Puga de la Cruz, & R. Torres Monzon, "REDES NEURONALES ARTIFICIALES PARA PRONOSTICAR EL RENDIMIENTO ACADÉMICO DE ALUMNOS DE INGENIERÍA DE SISTEMAS E INFORMÁTICA DE LA UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA". Repositorio Institucional Digital UNAP, Amazonas, 2023

[28] A. Oviedo Carrascal, & J. Jiménez Giraldo, "Minería de datos educativos: Análisis del desempeño de estudiantes de ingeniería en las pruebas saber-pro. Dialnet, 1-13. 2019

[29] A. Páez Rico, & N. Ramírez Gaytán, “Predicción del rendimiento académico utilizando las primeras actividades académicas de estudiantes universitarios y técnicas de aprendizaje automático”. (31 de agosto de 2022)

[30] O. Molina & D. Fuentes Cancel, “¿Se logra predecir el rendimiento académico? Un análisis desde latecnología educativa”. 363-375. (15 de setiembre de 2021). https://institucional.us.es/revistas/fuente/23_3/14278.pdf

[31] I. Méndez, A. Ramírez Reyes, & R. Mora-Gutiérrez, “Aprendizaje automático aplicado en física:Una revisión de la literatura científica”. 803–816. (2020)

[32] H. Packard, “Modelo de aprendizaje automático”. (23 de Agosto de 2023). https://www.hpe.com/lamerica/es/what-is/ml-models.html

[33] J. Gamboa Unsihuay & J. Salinas Flores, “Predicción De La Situación Académica En Alumnos De Pregrado Usando Algoritmos De Machine Learning” . 1(27). ,(31 de enero de 2022).

[34] L. E. Contreras, & H. J. Fuentes, “Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático”. 13(5). (2020)

Anexos

Anexo 1: Matriz de consistencia

<table>
<tr><td colspan="6">TITULO: Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete</td></tr>
<tr><td colspan="6">AUTORES: JEANICE DANIELA CHUMPITAZ LUNA // ANEL CECILIA HINOJOSA ROJAS</td></tr>
<tr><td rowspan="2">Problema</td><td rowspan="2">Objetivos</td><td rowspan="2">Hipótesis</td><td colspan="3">Variables e Indicadores</td></tr>
<tr><td colspan="3">Variable independiente: Sistema de predicción</td></tr>
<tr><td rowspan="3">¿Cómo predice un sistema el rendimiento académico de los estudiantes de la institución educativa Mis Primeros Pasos, Cañete?</td><td rowspan="3">Determinar la predicción del rendimiento académico de los estudiantes de la institución educativa Mis Primeros Pasos, Cañete.</td><td rowspan="3">La aplicación de minería de datos ayuda a predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.</td><td colspan="3">Variable dependiente: Rendimiento Académico</td></tr>
<tr><td>Dimensiones</td><td>Indicadores</td><td>Escala</td></tr>
<tr><td rowspan="2">Nivel Económico</td><td>Tipo de colegio</td><td rowspan="2">De razón</td></tr>
<tr><td>¿Qué indicadores de impacto pueden ayudar a determinar los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete?</td><td>Determinar los indicadores de impacto que ayuda a reconocer los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete.</td><td>Los factores encontrados influyen significativamente en la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.</td><td>Escala de pensión</td></tr>
<tr><td rowspan="3">Nivel demográfico</td><td>Edad</td><td rowspan="3">De razón</td></tr>
<tr><td rowspan="2">¿Cuál es el algoritmo óptimo para predecir el rendimiento académico de los estudiantes de la Institución Educativa</td><td rowspan="2">Determinar el algoritmo óptimo que genere un modelo para predecir el rendimiento académico de</td><td rowspan="2">Se puede determinar el algoritmo óptimo que genera un modelo predictivo en el rendimiento</td><td>Género</td></tr>
<tr><td>Residencia</td></tr>
<tr><td>Rendimiento</td><td>Calificación final al terminar el bimestre</td><td>De razón</td></tr>
</table>

<table>
<tr><td>Mis Primeros Pasos, Cañete?</td><td>los estudiantes de la Institución Educativas Mis Primeros Pasos, Cañete.</td><td>académico de los estudiantes de la</td><td></td><td>Asistencia por periodo
Curso</td><td></td></tr>
<tr><td></td><td></td><td>Institución Educativa Mis Primeros Pasos, Cañete.</td><td colspan="3">Metodología</td></tr>
<tr><td></td><td></td><td></td><td>Tipo: Aplicada</td><td>Población: 300 estudiantes</td><td>Técnica: Análisis documental</td></tr>
<tr><td rowspan="3">¿Como la predicción de rendimiento académico ayuda a detectar a los estudiantes que necesitan aprendizaje personalizado en la Institución Educativa Mis Primeros Pasos, Cañete?</td><td rowspan="3">Determinar una predicción al rendimiento académico que ayudará con el aprendizaje personalizado a los estudiantes de la Institución Educativa Mis Primeros Pasos.</td><td rowspan="3">La predicción de rendimiento académico detectó el aprendizaje personalizado que necesitan algunos estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.</td><td>Enfoque: Cuantitativo</td><td>Muestra: 2110 registros de notas</td><td>Instrumento: Registro de datos Históricos</td></tr>
<tr><td>Nivel: Explicativo</td><td colspan="2">Muestreo: No probabilístico por conveniencia</td></tr>
<tr><td colspan="3">Método: Teórico - analítico - sistémico, Teórico lógico - histórico, Teórico por modelación</td></tr>
<tr><td colspan="3">Diseño: No – Experimental</td></tr>
</table>

Anexo 2: Instrumento de recolección de datos

FICHA DE REGISTRO DE NOTAS

ALUMNO	TIPO DE COLEGIO	ESCALA DE PENSIÓN	EDAD	SEXO	RESIDENCIA	ARITMETICA	RAZONAMIENTO MATEMATICO	GEOMETRIA	ALGEBRA	TRIGONOMETRIA

FICHA DE ASISTENCIA

AÑO	PERIODO	CODIGO_ALUMNO	PRESENTE	TARDANZA INJUSTIFICADA	TARDANZA JUSTIFICADA	INASISTENCIA INJUSTIFICADA	INASISTENCIA JUSTIFICADA

Anexo 3: Validación de Instrumentos de Investigación

Validación del Experto 1

UNIVERSIDAD NACIONAL DE CAÑETE

VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN POR CRITERIO DE JUECES

I. DATOS GENERALES

1.1 Apellidos y nombres del juez : ....PH.D. DULIO OSEDA GAGO.....
1.2 Cargo e institución donde labora : ...DIRECTOR DEL INSTITUTO DE INVESTIGAICON UNDC..
1.3 Nombre del instrumento evaluado: ... FICHA DE REGISTRO DE DATOS.....
1.4. Autor (es) del instrumento : ... CHUMPITAZ LUNA JEANICE DANIELA / HINOJOSA ROJAS ANEL CECILIA

II. ASPECTO DE LA VALIDACIÓN

INDICADORES	CRITERIOS	Deficiente 1	Baja 2	Regular 3	Buena 4	Muy buena 5
1.CLARIDAD	Está formulado con lenguaje apropiado y comprensible					X
2.OBJETIVIDAD	Permite medir hechos observables.					x
3.ACTUALIDAD	Adecuado al avance de la ciencia y tecnología.					X
4.ORGANIZACIÓN	Presentación sistemática y ordenada.					x
5.SUFICIENCIA	Comprende aspectos de las variables en cantidad y calidad suficiente.					X
6.PERTINENCIA	Permite conseguir datos de acuerdo a los objetivos planteados.					X
7.CONSISTENCIA	Pretende conseguir datos basados en teorías o modelos teóricos.					X
8.COHERENCIA	Entre variables, indicadores y los ítems.					x
9.METODOLOGÍA	La metodología responde al propósito de la investigación.					X
10.APLICACIÓN	Los datos permiten un tratamiento estadístico pertinente y adecuado.					X
		↓	↓	↓	↓	↓
CONTEO TOTAL DE MARCAS (realice el conteo en cada una de las categorías de la escala)		A	B	C	D	50 E

$$Coeficiente\ de\ Validez = \frac{A+B+C+D+E}{50} = \frac{50}{50} = 1$$

III. CALIFICACIÓN GLOBAL (Ubique el coeficiente de validez obtenido en el intervalo respectivo y marque con un aspa en el círculo asociado)

Niveles	Condición	Intervalos
Desaprobado	( )	[0,00 – 0,50]
Observado	( )	[0,56 – 0,75]
Aprobado	(X)	[0,76 – 1,00]

IV. OPINIÓN DE APLICABILIDAD

.....APLICAR.....
.....

Lugar y fecha: ...Cañete, 24-05-2024.....

Firma del juez

FIRMA DIGITAL

Firmado digitalmente por: OSEDA GAGO DULIO
Motivo: Soy el autor del documento
Fecha: 24/05/2024 23:59:30-0500

Validación del Experto 2

UNIVERSIDAD NACIONAL DE CAÑETE

VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN POR CRITERIO DE JUECES

I. DATOS GENERALES

1.1 Apellidos y nombres del juez : ANGOMA ASTUCURI, MIRIAM.....
1.2 Cargo e institución donde labora : Universidad Nacional de Cañete.....
1.3 Nombre del instrumento evaluado: Ficha de registro del indicador 1.....
1.4. Autor (es) del instrumento : Chumpitaz Luna Jeanice Daniela / Hinojosa Rojas Anel C.

II. ASPECTO DE LA VALIDACIÓN

INDICADORES	CRITERIOS	Deficiente 1	Baja 2	Regular 3	Buena 4	Muy buena 5
1.CLARIDAD	Está formulado con lenguaje apropiado y comprensible				X	
2.OBJETIVIDAD	Permite medir hechos observables.					X
3.ACTUALIDAD	Adecuado al avance de la ciencia y tecnología.				X	
4.ORGANIZACIÓN	Presentación sistemática y ordenada.				X	
5.SUFICIENCIA	Comprende aspectos de las variables en cantidad y calidad suficiente.				X	
6.PERTINENCIA	Permite conseguir datos de acuerdo a los objetivos planteados.					X
7.CONSISTENCIA	Pretende conseguir datos basados en teorías o modelos teóricos.				X	
8.COHERENCIA	Entre variables, indicadores y los ítems.				X	
9.METODOLOGÍA	La metodología responde al propósito de la investigación.				X	
10.APLICACIÓN	Los datos permiten un tratamiento estadístico pertinente y adecuado.				X	
		↓	↓	↓	↓	↓
CONTEO TOTAL DE MARCAS (realice el conteo en cada una de las categorías de la escala)		0 A	0 B	0 C	32 D	10 E

$$Coeficiente\ de\ Validez = \frac{A+B+C+D+E}{5} = \frac{42}{5*10} = 0.84$$

III. CALIFICACIÓN GLOBAL (Ubique el coeficiente de validez obtenido en el intervalo respectivo y marque con un aspa en el círculo asociado)

Niveles	Condición	Intervalos
Desaprobado	( )	[0,00 – 0,50]
Observado	( )	[0,56 – 0,75]
Aprobado	(X)	[0,76 – 1,00]

IV. OPINIÓN DE APLICABILIDAD

El instrumento es aplicable para obtener los datos requeridos en la investigación.

Lugar y fecha: 20.....San Vicente de Cañete 2024.....

FIRMA DIGITAL

Firmado digitalmente por: ANGOMA ASTUCURI MIRIAM FIR 20044824 hard
Motivo: Soy el autor del documento
Fecha: 23/05/2024 14:52:08-0500

Firma del juez

Validación del Experto 3

UNIVERSIDAD NACIONAL DE CAÑETE

VALIDACIÓN DEL INSTRUMENTO DE INVESTIGACIÓN POR CRITERIO DE JUECES

I. DATOS GENERALES

1.1 Apellidos y nombres del juez : MG. AMANDA DURÁN CARHUAMACA
1.2 Cargo e institución donde labora : DOCENTE DE LA UNDC
1.3 Nombre del instrumento evaluado: FICHA DE REGISTRO DE DATOS
1.4. Autor (es) del instrumento : CHUMPITAZ LUNA JEANICE DANIELA / HINOJOSA ROJAS ANEL CECILIA

II. ASPECTO DE LA VALIDACIÓN

INDICADORES	CRITERIOS	Deficiente 1	Baja 2	Regular 3	Buena 4	Muy buena 5
1.CLARIDAD	Está formulado con lenguaje apropiado y comprensible					X
2.OBJETIVIDAD	Permite medir hechos observables.					X
3.ACTUALIDAD	Adecuado al avance de la ciencia y tecnología.					X
4.ORGANIZACIÓN	Presentación sistemática y ordenada.					X
5.SUFICIENCIA	Comprende aspectos de las variables en cantidad y calidad suficiente.					X
6.PERTINENCIA	Permite conseguir datos de acuerdo a los objetivos planteados.					X
7.CONSISTENCIA	Pretende conseguir datos basados en teorías o modelos teóricos.					X
8.COHERENCIA	Entre variables, indicadores y los ítems.					X
9.METODOLOGÍA	La metodología responde al propósito de la investigación.					X
10.APLICACIÓN	Los datos permiten un tratamiento estadístico pertinente y adecuado.					X
		↓	↓	↓	↓	↓
CONTEO TOTAL DE MARCAS (realice el conteo en cada una de las categorías de la escala)		A	B	C	D	15 E

$$Coeficiente\ de\ Validez = \frac{A + B + C + D + E}{50} = \frac{50}{50} = 1$$

III. CALIFICACIÓN GLOBAL (Ubique el coeficiente de validez obtenido en el intervalo respectivo y marque con un aspa en el círculo asociado)

Niveles	Condición	Intervalos
Desaprobado	( )	[0,00 – 0,50]
Observado	( )	[0,56 – 0,75]
Aprobado	(X)	[0,76 – 1,00]

IV. OPINIÓN DE APLICABILIDAD

APLICAR

Lugar y fecha: CAÑETE, 27/05/2024

Firma del juez

FIRMA DIGITAL

Firmado digitalmente por: DURAN CARHUAMACA AMANDA FIR 20114878 hard
Motivo: En señal de conformidad
Fecha: 27/05/2024 09:24:13-0500

Anexo 4: Constancia de grados y títulos de los validadores (SUNEDU)

Experto 1

REGISTRO NACIONAL DE
GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES

Graduado	Grado o Título	Institución
OSEDA GAGO, DULIO **DNI 20044737**	DOCTOR EN PSICOLOGIA EDUCACIONAL Y TUTORIAL Fecha de diploma: 27/08/2012 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	**UNIVERSIDAD NACIONAL DE EDUCACIÓN ENRIQUE GUZMÁN Y VALLE** *PERU*
OSEDA GAGO, DULIO **DNI 20044737**	MAGISTER EN EDUCACION DOCENCIA EN EDUCACION SUPERIOR Fecha de diploma: 12/01/2009 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	**UNIVERSIDAD PERUANA LOS ANDES** *PERU*
OSEDA GAGO, DULIO **DNI 20044737**	MAGISTER EN ADMINISTRACION DE LA EDUCACION Fecha de diploma: 28/05/2014 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	**UNIVERSIDAD PRIVADA CÉSAR VALLEJO** *PERU*
	INGENIERO DE SISTEMAS Y	

Experto 2

REGISTRO NACIONAL DE
GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES

Graduado	Grado o Título	Institución
ANGOMA ASTUCURI, MIRIAM DNI 20044924	INGENIERA DE SISTEMAS Y COMPUTACION INGENIERIA DE SISTEMAS Y COMPUTACION Fecha de diploma: 17/03/2008 Modalidad de estudios: -	UNIVERSIDAD PERUANA LOS ANDES *PERU*
ANGOMA ASTUCURI, MIRIAM DNI 20044924	BACHILLER EN INGENIERIA DE SISTEMAS Y COMPUTACION Fecha de diploma: 22/03/2004 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	UNIVERSIDAD PERUANA LOS ANDES *PERU*
ANGOMA ASTUCURI, MIRIAM DNI 20044924	MAGÍSTER EN ADMINISTRACIÓN GESTIÓN DE PROYECTOS Fecha de diploma: 14/10/15 Modalidad de estudios: PRESENCIAL Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	UNIVERSIDAD NACIONAL DEL CENTRO DEL PERÚ *PERU*
ANGOMA ASTUCURI, MIRIAM DNI 20044924	DOCTORA EN SISTEMAS DE INGENIERÍA Fecha de diploma: 14/02/17 Modalidad de estudios: PRESENCIAL	UNIVERSIDAD NACIONAL DEL CENTRO DEL PERÚ *PERU*

Experto 3

REGISTRO NACIONAL DE
GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES

Graduado	Grado o Título	Institución
DURAN CARHUAMACA, AMANDA **DNI 20114878**	**MAGISTER EN INFORMATICA** Fecha de diploma: 02/07/2003 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	**PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ** *PERU*
DURAN CARHUAMACA, AMANDA **DNI 20114878**	**INGENIERA DE SISTEMAS** Fecha de diploma: 21/07/2004 Modalidad de estudios: -	**UNIVERSIDAD NACIONAL DEL CENTRO DEL PERÚ** *PERU*
DURAN CARHUAMACA, AMANDA **DNI 20114878**	**BACHILLER EN INGENIERIA DE SISTEMAS** Fecha de diploma: 12/01/2000 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	**UNIVERSIDAD NACIONAL DEL CENTRO DEL PERÚ** *PERU*

Anexo 5: Bases de datos

REGISTRO DE NOTAS

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETR
ARJE6202	PRIVADO	?	12	M	MALA	15	14	12	13	14
CALAN6149	PRIVADO	?	12	F	MALA	16	17	16	14	15
CODA6066	PRIVADO	?	12	M	MALA	20	17	19	16	19
DINA6070	PRIVADO	?	12	F	MALA	20	19	18	18	18
LALU6066	PRIVADO	?	12	F	MALA	16	16	15	13	11
MAAN6070	PRIVADO	?	12	M	MALA	14	14	12	11	11
PUJO6203	PRIVADO	?	12	M	MALA	15	15	15	14	13
REJO6040	PRIVADO	?	12	M	MALA	11	13	13	13	10
RIYHA6138	PRIVADO	?	12	F	MALA	19	17	18	16	14
RIMA6070	PRIVADO	?	12	F	MALA	11	12	11	10	10
SAIK6063	PRIVADO	?	12	M	MALA	16	17	14	15	15
SOMI6151	PRIVADO	?	12	F	MALA	14	14	14	12	12
VAJO6203	PRIVADO	?	12	M	MALA	12	12	13	11	11
VEPA6138	PRIVADO	?	12	M	MALA	20	19	20	17	18
VILA6007	PRIVADO	?	12	F	MALA	11	10	11	09	09

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETR
ARJE6202	PRIVADO	?	12	M	MALA	15	14	12	13	14
CALAN6149	PRIVADO	?	12	F	MALA	16	17	16	14	15
CODA6066	PRIVADO	?	12	M	MALA	20	17	19	16	19
DINA6070	PRIVADO	?	12	F	MALA	20	19	18	18	18
LALU6066	PRIVADO	?	12	F	MALA	16	16	15	13	11
MAAN6070	PRIVADO	?	12	M	MALA	14	14	12	11	11
PUJO6203	PRIVADO	?	12	M	MALA	15	15	15	14	13
REJO6040	PRIVADO	?	12	M	MALA	11	13	13	13	10
RIYHA6138	PRIVADO	?	12	F	MALA	19	17	18	16	14
RIMA6070	PRIVADO	?	12	F	MALA	11	12	11	10	10
SAIK6063	PRIVADO	?	12	M	MALA	16	17	14	15	15
SOMI6151	PRIVADO	?	12	F	MALA	14	14	14	12	12
VAJO6203	PRIVADO	?	12	M	MALA	12	12	13	11	11
VEPA6138	PRIVADO	?	12	M	MALA	20	19	20	17	18
VILA6007	PRIVADO	?	12	F	MALA	11	10	11	09	09

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍ
ARJE6202	PRIVADO	?	12	M	MALA	14	14	14	12	13
CALAN6149	PRIVADO	?	12	F	MALA	17	18	18	15	16
CODA6066	PRIVADO	?	12	M	MALA	19	20	19	19	19
DINA6070	PRIVADO	?	12	F	MALA	18	18	19	18	18
LALU6066	PRIVADO	?	12	F	MALA	14	13	17	14	14
MAAN6070	PRIVADO	?	12	M	MALA	16	14	16	13	14
PUJO6203	PRIVADO	?	12	M	MALA	16	14	17	16	15
REJO6040	PRIVADO	?	12	M	MALA	12	13	13	11	12
RIYHA6138	PRIVADO	?	12	F	MALA	16	18	19	18	15
RIMA6070	PRIVADO	?	12	F	MALA	14	13	12	12	11
SAIK6063	PRIVADO	?	12	M	MALA	16	19	19	19	18
SOMI6151	PRIVADO	?	12	F	MALA	16	14	15	16	13
VAJO6203	PRIVADO	?	12	M	MALA	12	12	15	11	10
VEPA6138	PRIVADO	?	12	M	MALA	18	19	19	19	17
VILA6007	PRIVADO	?	12	F	MALA	10	10	10	10	10

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
ACAL6238	PRIVADO	?	12	M	MALA	16	17	17	19	18
ASZA6199	PRIVADO	?	12	F	MALA	13	15	13	15	15
BALVI6038	PRIVADO	?	12	M	MALA	17	20	19	19	20
CARES6196	PRIVADO	?	12	F	MALA	14	16	15	15	13
CHUASH2020	PRIVADO	?	12	F	MALA	12	15	13	13	13
DIJA6235	PRIVADO	?	12	F	MALA	13	13	14	14	12
FELE6166	PRIVADO	?	12	M	MALA	14	15	14	15	17
FRALE6038	PRIVADO	?	12	F	MALA	12	13	13	14	12
HUASHAN74	PRIVADO	?	12	F	MALA	12	12	12	11	12
LOZHA6198	PRIVADO	?	12	F	MALA	16	19	18	18	18
MOSE6198	PRIVADO	?	12	M	MALA	13	14	13	14	13
TEJU7384	PRIVADO	?	12	M	MALA	11	12	13	15	12
TRUJE7355	PRIVADO	?	12	M	MALA	15	18	17	17	15
ARJE6202	PRIVADO	?	13	M	MALA	13	14	15	14	14
CALAN6149	PRIVADO	?	13	F	MALA	15	15	14	16	17
CHUKA6165	PRIVADO	?	13	F	MALA	14	15	16	15	15
CODA6066	PRIVADO	?	13	M	MALA	19	17	19	18	20
DINA6070	PRIVADO	?	13	F	MALA	17	18	16	17	20
LALU6066	PRIVADO	?	13	F	MALA	12	15	14	12	15
MAAN6070	PRIVADO	?	13	M	MALA	11	13	12	13	13
NAKA6035	PRIVADO	?	13	F	MALA	14	15	15	16	16
PUJO6203	PRIVADO	?	13	M	MALA	17	17	16	16	17
REJO6040	PRIVADO	?	13	M	MALA	12	15	13	14	16
RIYHA6138	PRIVADO	?	13	F	MALA	17	17	18	18	18
RIMA6070	PRIVADO	?	13	F	MALA	11	14	12	12	12
SAIK6063	PRIVADO	?	13	M	MALA	16	17	16	18	18
SISE2020	PRIVADO	?	13	M	MALA	12	14	14	15	13
SOMI6151	PRIVADO	?	13	F	MALA	13	14	12	13	14
VEPA6138	PRIVADO	?	13	M	MALA	19	18	19	17	19

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍ
ACAL6238	PRIVADO	?	12	M	MALA	16.5	17	15	16	18
ASZA6199	PRIVADO	?	12	F	MALA	16	14	13	15	15
BALVI6038	PRIVADO	?	12	M	MALA	20	20	20	20	19
CARES6196	PRIVADO	?	12	F	MALA	13.5	15	14	14	16
CHUASH2020	PRIVADO	?	12	F	MALA	11.5	14	13	12	11
DIJA6235	PRIVADO	?	12	F	MALA	11.5	12	13	12	14
FELE6166	PRIVADO	?	12	M	MALA	18	16	16	15	18
FRALE6038	PRIVADO	?	12	F	MALA	12.5	15	12	13	14
GUAG6238	PRIVADO	?	12	M	MALA	11.5	14	12	11	14
HUASHAN74	PRIVADO	?	12	F	MALA	11	11	11	10	13
LOZHA6198	PRIVADO	?	12	F	MALA	17.5	19	16	17	20
MOSE6198	PRIVADO	?	12	M	MALA	11.5	12	12	11	13
TEJU7384	PRIVADO	?	12	M	MALA	12.5	12	12	11	11
TRUJE7355	PRIVADO	?	12	M	MALA	19	16	15	15	16
ARJE6202	PRIVADO	?	13	M	MALA	13	13	13	12	13
CALAN6149	PRIVADO	?	13	F	MALA	14	15	17	13	13
CHUKA6165	PRIVADO	?	13	F	MALA	14	16	16	11	13
CODA6066	PRIVADO	?	13	M	MALA	20	18	20	18	20
DINA6070	PRIVADO	?	13	F	MALA	19.5	20	19	18	18
LALU6066	PRIVADO	?	13	F	MALA	12.5	14	14	12	12
MAAN6070	PRIVADO	?	13	M	MALA	12	13	14	11	12
NAKA6035	PRIVADO	?	13	F	MALA	15.5	16	15	11	13
PUJO6203	PRIVADO	?	13	M	MALA	17.5	19	17	16	18
REJO6040	PRIVADO	?	13	M	MALA	12.5	16	13	11	14
RIYHA6138	PRIVADO	?	13	F	MALA	19	19	19	19	18
RIMA6070	PRIVADO	?	13	F	MALA	12.5	12	13	12	12
SAIK6063	PRIVADO	?	13	M	MALA	19	18	18	18	16
SISE2020	PRIVADO	?	13	M	MALA	14.5	15	15	11	14
SOMI6151	PRIVADO	?	13	F	MALA	14	12	14	14	12
VEPA6138	PRIVADO	?	13	M	MALA	20	18	19	19	18

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
ACAL6238	PRIVADO	?	12	M	MALA	16	19	16	15	19
ASZA6199	PRIVADO	?	12	F	MALA	13	18	15	12	14
BALVI6038	PRIVADO	?	12	M	MALA	19	19	20	19	20
BRANI6199	PRIVADO	?	12	F	MALA	12	12	11	11	12
CARES6196	PRIVADO	?	12	F	MALA	15	17	16	13	15
CHUASH2020	PRIVADO	?	12	F	MALA	13	12	13	12	11
DIJA6235	PRIVADO	?	12	F	MALA	14	13	13	12	11
FELE6166	PRIVADO	?	12	M	MALA	16	20	18	15	17
FRALE6038	PRIVADO	?	12	F	MALA	12	15	14	12	12
GUAG6238	PRIVADO	?	12	M	MALA	11	12	11	10	11
HUASHAN74	PRIVADO	?	12	F	MALA	10	11	11	11	12
LOZHA6198	PRIVADO	?	12	F	MALA	18	20	20	18	20
MOSE6198	PRIVADO	?	12	M	MALA	11	13	11	10	10
TEJU7384	PRIVADO	?	12	M	MALA	11	14	12	11	11
TRUJE7355	PRIVADO	?	12	M	MALA	16	18	17	16	19
ARJE6202	PRIVADO	?	13	M	MALA	16	16	15	12	11
CALAN6149	PRIVADO	?	13	F	MALA	13	15	14	11	12
CHUKA6165	PRIVADO	?	13	F	MALA	14	18	16	12	13
CODA6066	PRIVADO	?	13	M	MALA	18	20	19	18	19
DINA6070	PRIVADO	?	13	F	MALA	18	20	20	17	18
LALU6066	PRIVADO	?	13	F	MALA	13	15	16	12	14
MAAN6070	PRIVADO	?	13	M	MALA	10	13	13	10	10
NAKA6035	PRIVADO	?	13	F	MALA	14	17	15	13	11
PUJO6203	PRIVADO	?	13	M	MALA	18	20	19	16	14
REJO6040	PRIVADO	?	13	M	MALA	14	13	12	11	11
RIYHA6138	PRIVADO	?	13	F	MALA	18	19	19	18	18
RIMA6070	PRIVADO	?	13	F	MALA	12	12	13	11	11
SAIK6063	PRIVADO	?	13	M	MALA	18	19	19	18	18
SISE2020	PRIVADO	?	13	M	MALA	16	16	15	12	14
SOMI6151	PRIVADO	?	13	F	MALA	11	13	12	12	11
VEPA6138	PRIVADO	?	13	M	MALA	19	20	19	18	19

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
ACAL6238	PRIVADO	?	12	M	MALA	16	19	17	17	15
ASZA6199	PRIVADO	?	12	F	MALA	11	15	13	14	13
BALVI6038	PRIVADO	?	12	M	MALA	20	20	20	20	20
BRANI6199	PRIVADO	?	12	F	MALA	11	12	10	11	10
CARES6196	PRIVADO	?	12	F	MALA	17	16	17	15	15
CHUASH2020	PRIVADO	?	12	F	MALA	14	13	12	13	14
DIJA6235	PRIVADO	?	12	F	MALA	12	14	11	14	12
FELE6166	PRIVADO	?	12	M	MALA	18	17	17	17	16
FRALE6038	PRIVADO	?	12	F	MALA	13	15	13	13	13
GUAG6238	PRIVADO	?	12	M	MALA	11	13	12	11	12
HUASHAN74	PRIVADO	?	12	F	MALA	11	13	11	11	11
LOZHA6198	PRIVADO	?	12	F	MALA	20	20	19	19	19
MOSE6198	PRIVADO	?	12	M	MALA	11	13	11	12	11
TEJU7384	PRIVADO	?	12	M	MALA	13	14	13	13	13
TRUJE7355	PRIVADO	?	12	M	MALA	18	18	17	17	19
ARJE6202	PRIVADO	?	13	M	MALA	13	16	14	13	12
CALAN6149	PRIVADO	?	13	F	MALA	15	16	12	11	12
CHUKA6165	PRIVADO	?	13	F	MALA	15	16	14	14	14
CODA6066	PRIVADO	?	13	M	MALA	19	20	20	18	20
DINA6070	PRIVADO	?	13	F	MALA	20	19	17	19	20
LALU6066	PRIVADO	?	13	F	MALA	13	16	13	12	12
MAAN6070	PRIVADO	?	13	M	MALA	11	13	11	12	10
NAKA6035	PRIVADO	?	13	F	MALA	14	14	12	12	11
PUJO6203	PRIVADO	?	13	M	MALA	17	19	18	16	17
REJO6040	PRIVADO	?	13	M	MALA	12	13	12	13	12
RIYHA6138	PRIVADO	?	13	F	MALA	19	19	19	17	20
SAIK6063	PRIVADO	?	13	M	MALA	18	20	19	18	20
SISE2020	PRIVADO	?	13	M	MALA	16	16	14	15	12
SOMI6151	PRIVADO	?	13	F	MALA	11	13	11	11	11
VEPA6138	PRIVADO	?	13	M	MALA	20	19	19	18	20

ALUMNO	Tipo de col	Escala de p	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
CUWA6273	PRIVADO	270	12	M	MALA	13	14	12	14	16
FRAKA6230	PRIVADO	270	12	F	MALA	13	13	12	13	13
GODAN7479	PRIVADO	270	12	F	MALA	18	16	15	16	16
HUANN6249	PRIVADO	270	12	F	MALA	17	15	15	16	18
HUAJO6196	PRIVADO	270	12	M	MALA	11	12	11	10	10
LANE6273	PRIVADO	270	12	M	MALA	12	13	11	11	11
SALU6224	PRIVADO	270	12	M	MALA	15	15	14	18	17
TOGRE6249	PRIVADO	270	12	F	MALA	13	13	12	14	17
VIYA6257	PRIVADO	270	12	F	MALA	15	17	13	15	15
ACAL6238	PRIVADO	270	13	M	MALA	17	17	15	14	15
ASZA6199	PRIVADO	270	13	F	MALA	16	12	14	13	14
BALVI6038	PRIVADO	270	13	M	MALA	19	19	19	19	20
BRANI6199	PRIVADO	270	13	F	MALA	11	12	11	11	11
CARES6196	PRIVADO	270	13	F	MALA	15	16	14	14	18
DIJA6235	PRIVADO	270	13	M	MALA	13	12	11	10	12
FELE6166	PRIVADO	270	13	F	MALA	15	14	15	14	18
FRALE6038	PRIVADO	270	13	F	MALA	13	13	13	12	13
GAGI6198	PRIVADO	270	13	M	MALA	14	13	13	13	12
GUAGU6238	PRIVADO	270	13	M	MALA	12	11	10	10	10
HERI6216	PRIVADO	270	13	M	MALA	11	11	11	11	12
HUASHAN74	PRIVADO	270	13	F	MALA	11	10	10	09	10
LOZHA6198	PRIVADO	270	13	F	MALA	19	19	18	17	19
MOSE6198	PRIVADO	270	13	M	MALA	12	11	11	11	11
NADA6207	PRIVADO	270	13	M	MALA	11	12	12	11	11
PEAR2020	PRIVADO	270	13	F	MALA	18	16	19	19	19
TEMO7384	PRIVADO	270	13	M	MALA	11	11	12	11	13
TRUJE7355	PRIVADO	270	13	M	MALA	17	15	16	16	18
VAYO6238	PRIVADO	270	13	F	MALA	18	18	17	17	19
VIRO6235	PRIVADO	270	13	F	MALA	13	13	13	11	12
ARJE6202	PRIVADO	270	14	M	MALA	14	13	13	13	13
CAAN6149	PRIVADO	270	14	F	MALA	12	12	13	11	13
CECI6066	PRIVADO	270	14	F	MALA	13	11	13	11	12
CHUKA6165	PRIVADO	270	14	F	MALA	16	15	14	14	13
CODA6066	PRIVADO	270	14	M	MALA	18	19	20	19	19
DINA6070	PRIVADO	270	14	F	MALA	19	18	19	18	18
LALU6066	PRIVADO	270	14	F	MALA	15	13	13	13	13
MABE7298	PRIVADO	270	14	F	MALA	11	13	13	12	12
MAAN6070	PRIVADO	270	14	M	MALA	10	11	11	11	10
NAKA6035	PRIVADO	270	14	F	MALA	15	16	16	14	14
PUJO6203	PRIVADO	270	14	M	MALA	19	17	15	18	17
REJO6040	PRIVADO	270	14	M	MALA	12	12	16	12	13
RIYHA6138	PRIVADO	270	14	F	MALA	17	16	20	18	19
RIMA6070	PRIVADO	270	14	F	MALA	15	12	13	12	12
SAIK6063	PRIVADO	270	14	M	MALA	18	16	18	17	17
SISE2020	PRIVADO	270	14	M	MALA	14	14	14	14	13
SOLU7135	PRIVADO	270	14	F	MALA	11	12	11	11	12
SOMI6151	PRIVADO	270	14	F	MALA	13	12	14	11	12
VEPA6138	PRIVADO	270	14	M	MALA	19	17	19	18	19

ALUMNO	Tipo de coleg	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIEN	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
CUWA6273	PRIVADO	270	12	M	MALA	11	17	13	12	12
FRAKA6230	PRIVADO	270	12	F	MALA	13	18	14	12	12
GODAN7479	PRIVADO	270	12	F	MALA	14	16	15	14	14
HUANN6249	PRIVADO	270	12	F	MALA	13	16	15	16	15
HUAJO6196	PRIVADO	270	12	M	MALA	09	12	10	09	9
LANE6273	PRIVADO	270	12	M	MALA	11	15	12	11	11
PABRI6217	PRIVADO	270	12	F	MALA	10	18	14	15	14
ROED7530	PRIVADO	270	12	M	MALA	12	15	13	13	12
SALU6224	PRIVADO	270	12	M	MALA	12	18	14	12	13
TOGRE6249	PRIVADO	270	12	F	MALA	12	16	13	12	13
VIYA6257	PRIVADO	270	12	F	MALA	11	18	17	13	12
ACAL6238	PRIVADO	270	13	M	MALA	17	18	17	16	17
ASZA6199	PRIVADO	270	13	F	MALA	15	15	16	12	14
BALVI6038	PRIVADO	270	13	M	MALA	19	19	18	19	20
BRANI6199	PRIVADO	270	13	F	MALA	09	10	10	10	10
CARES6196	PRIVADO	270	13	F	MALA	18	17	15	15	17
DIJA6235	PRIVADO	270	13	M	MALA	15	13	12	12	14
FELE6166	PRIVADO	270	13	F	MALA	18	19	17	16	19
FRALE6038	PRIVADO	270	13	F	MALA	13	15	12	11	13
GAGI6198	PRIVADO	270	13	M	MALA	14	13	14	12	13
GUAGU6238	PRIVADO	270	13	M	MALA	14	13	12	11	12
HERI6216	PRIVADO	270	13	M	MALA	14	15	13	11	12
HUASHAN740	PRIVADO	270	13	F	MALA	09	09	10	10	10
LOZHA6198	PRIVADO	270	13	F	MALA	18	17	17	16	19
MOSE6198	PRIVADO	270	13	M	MALA	14	14	13	12	13
NADA6207	PRIVADO	270	13	M	MALA	10	11	10	11	12
PEAR2020	PRIVADO	270	13	F	MALA	19	19	19	18	19
TEMO7384	PRIVADO	270	13	M	MALA	15	14	13	11	12
TRUJE7355	PRIVADO	270	13	M	MALA	19	16	15	15	18
VAYO6238	PRIVADO	270	13	F	MALA	19	18	18	18	19
VIRO6235	PRIVADO	270	13	F	MALA	13	15	12	12	13
ARJE6202	PRIVADO	270	14	M	MALA	13	16	15	12	14
CAAN6149	PRIVADO	270	14	F	MALA	12	15	13	12	12
CECI6066	PRIVADO	270	14	F	MALA	10	11	12	10	11
CHUKA6165	PRIVADO	270	14	F	MALA	13	15	15	13	15
CODA6066	PRIVADO	270	14	M	MALA	19	18	19	17	20
DINA6070	PRIVADO	270	14	F	MALA	19	20	19	17	20
LALU6066	PRIVADO	270	14	F	MALA	12	14	14	11	14
MABE7298	PRIVADO	270	14	F	MALA	12	15	15	12	13
MAAN6070	PRIVADO	270	14	M	MALA	12	12	12	11	12
NAKA6035	PRIVADO	270	14	F	MALA	12	15	14	11	14
PUJO6203	PRIVADO	270	14	M	MALA	18	20	19	16	19
REJO6040	PRIVADO	270	14	M	MALA	12	15	12	11	12
RIYHA6138	PRIVADO	270	14	F	MALA	18	19	18	17	18
RIMA6070	PRIVADO	270	14	F	MALA	12	13	12	11	12
SAIK6063	PRIVADO	270	14	M	MALA	19	18	17	17	19
SISE2020	PRIVADO	270	14	M	MALA	16	18	17	13	15
SOLU7135	PRIVADO	270	14	F	MALA	12	15	12	10	11
SOMI6151	PRIVADO	270	14	F	MALA	11	13	12	10	11
VEPA6138	PRIVADO	270	14	M	MALA	19	20	20	17	20

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
CUWA6273	PRIVADO	270	12	M	MALA	13	11	13	12	14
FRAKA6230	PRIVADO	270	12	F	MALA	14	12	14	13	15
GODAN7479	PRIVADO	270	12	F	MALA	16	15	17	16	18
HUANN6249	PRIVADO	270	12	F	MALA	15	15	16	14	17
HUAJO6196	PRIVADO	270	12	M	MALA	12	11	11	11	12
LANE6273	PRIVADO	270	12	M	MALA	11	10	11	10	12
PABRI6217	PRIVADO	270	12	F	MALA	16	17	17	15	19
ROED7530	PRIVADO	270	12	M	MALA	14	14	14	12	18
SALU6224	PRIVADO	270	12	M	MALA	15	13	15	14	15
TEDA6250	PRIVADO	270	12	F	MALA	11	11	10	10	11
TOGRE6249	PRIVADO	270	12	F	MALA	13	12	12	12	15
VIYA6257	PRIVADO	270	12	F	MALA	14	14	13	13	13
ACAL6238	PRIVADO	270	13	M	MALA	15	16	18	16	15
ASZA6199	PRIVADO	270	13	F	MALA	13	14	15	13	14
BALVI6038	PRIVADO	270	13	M	MALA	17	19	20	17	20
BRANI6199	PRIVADO	270	13	F	MALA	10	09	12	09	10
CARES6196	PRIVADO	270	13	F	MALA	14	17	19	15	16
DIJA6235	PRIVADO	270	13	M	MALA	13	15	14	13	13
FELE6166	PRIVADO	270	13	F	MALA	14	17	15	14	15
FRALE6038	PRIVADO	270	13	F	MALA	11	12	12	11	12
GAGI6198	PRIVADO	270	13	M	MALA	12	13	13	12	13
GUAGU6238	PRIVADO	270	13	M	MALA	11	10	13	10	11
HERI6216	PRIVADO	270	13	M	MALA	11	11	12	12	12
LOZHA6198	PRIVADO	270	13	F	MALA	17	18	19	18	16
MOSE6198	PRIVADO	270	13	M	MALA	12	12	12	11	13
NADA6207	PRIVADO	270	13	M	MALA	10	12	10	10	11
PEAR2020	PRIVADO	270	13	F	MALA	15	16	18	17	17
TEMO7384	PRIVADO	270	13	M	MALA	11	12	12	11	13
TRUJE7355	PRIVADO	270	13	M	MALA	15	14	16	14	15
VAYO6238	PRIVADO	270	13	F	MALA	18	15	18	16	16
VIRO6235	PRIVADO	270	13	F	MALA	11	13	15	12	12
ARJE6202	PRIVADO	270	14	M	MALA	13	17	16	14	14
CAAN6149	PRIVADO	270	14	F	MALA	12	16	15	12	13
CECI6066	PRIVADO	270	14	F	MALA	11	12	12	11	11
CHUKA6165	PRIVADO	270	14	F	MALA	15	15	15	12	15
CHUBRI6070	PRIVADO	270	14	F	MALA	10	13	13	10	11
CODA6066	PRIVADO	270	14	M	MALA	18	20	20	19	20
DINA6070	PRIVADO	270	14	F	MALA	17	19	19	18	19
LALU6066	PRIVADO	270	14	F	MALA	13	16	16	13	15
MABE7298	PRIVADO	270	14	F	MALA	12	16	14	13	14
MAAN6070	PRIVADO	270	14	M	MALA	10	12	15	11	11
NAKA6035	PRIVADO	270	14	F	MALA	12	13	16	12	16
PUJO6203	PRIVADO	270	14	M	MALA	18	17	18	17	19
REJO6040	PRIVADO	270	14	M	MALA	12	13	13	11	13
RIYHA6138	PRIVADO	270	14	F	MALA	16	20	19	19	20
RIMA6070	PRIVADO	270	14	F	MALA	12	13	12	11	12
SAIK6063	PRIVADO	270	14	M	MALA	18	18	19	18	19
SISE2020	PRIVADO	270	14	M	MALA	14	18	18	15	16
SOLU7135	PRIVADO	270	14	F	MALA	10	14	14	11	12
SOMI6151	PRIVADO	270	14	F	MALA	11	14	12	11	12
VEPA6138	PRIVADO	270	14	M	MALA	19	19	20	19	20

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
CUWA6273	PRIVADO	270	12	M	MALA	12	12	12	14	12
FRAKA6230	PRIVADO	270	12	F	MALA	15	15	14	11	12
GODAN7479	PRIVADO	270	12	F	MALA	18	16	16	12	14
HUANN6249	PRIVADO	270	12	F	MALA	18	17	17	15	16
HUAJO6196	PRIVADO	270	12	M	MALA	13	12	12	11	12
LANE6273	PRIVADO	270	12	M	MALA	12	11	11	11	11
PABRI6217	PRIVADO	270	12	F	MALA	19	16	16	11	13
ROED7530	PRIVADO	270	12	M	MALA	18	12	13	10	11
SALU6224	PRIVADO	270	12	M	MALA	14	12	13	12	13
TEDA6250	PRIVADO	270	12	F	MALA	11	10	10	11	12
VIYA6257	PRIVADO	270	12	F	MALA	17	16	14	13	13
ACAL6238	PRIVADO	270	13	M	MALA	14	18	15	12	14
ASZA6199	PRIVADO	270	13	F	MALA	15	16	16	12	14
BALVI6038	PRIVADO	270	13	M	MALA	20	20	20	17	20
BRANI6199	PRIVADO	270	13	F	MALA	11	12	10	10	11
CARES6196	PRIVADO	270	13	F	MALA	16	17	17	14	16
DIJA6235	PRIVADO	270	13	M	MALA	13	14	15	11	14
FELE6166	PRIVADO	270	13	F	MALA	13	15	17	13	16
FRALE6038	PRIVADO	270	13	F	MALA	13	14	14	11	13
GAGI6198	PRIVADO	270	13	M	MALA	14	16	14	11	14
GUAGU6238	PRIVADO	270	13	M	MALA	13	14	13	10	14
HERI6216	PRIVADO	270	13	M	MALA	12	14	13	10	12
LOZHA6198	PRIVADO	270	13	F	MALA	19	20	19	15	18
MOSE6198	PRIVADO	270	13	M	MALA	11	13	12	10	12
NADA6207	PRIVADO	270	13	M	MALA	12	14	12	10	12
PEAR2020	PRIVADO	270	13	F	MALA	19	19	19	15	17
TEMO7384	PRIVADO	270	13	M	MALA	13	15	12	11	13
TRUJE7355	PRIVADO	270	13	M	MALA	15	18	17	13	15
VAYO6238	PRIVADO	270	13	F	MALA	19	20	19	16	20
VIRO6235	PRIVADO	270	13	F	MALA	13	15	15	12	13
ARJE6202	PRIVADO	270	14	M	MALA	14	19	18	14	14
CAAN6149	PRIVADO	270	14	F	MALA	13	14	14	13	14
CECI6066	PRIVADO	270	14	F	MALA	13	15	14	12	13
CHUKA6165	PRIVADO	270	14	F	MALA	15	19	15	14	16
CHUBRI6070	PRIVADO	270	14	F	MALA	12	15	14	12	12
CODA6066	PRIVADO	270	14	M	MALA	18	20	20	17	17
DINA6070	PRIVADO	270	14	F	MALA	18	20	20	19	18
LALU6066	PRIVADO	270	14	F	MALA	15	17	15	15	15
MABE7298	PRIVADO	270	14	F	MALA	15	17	15	14	14
MAAN6070	PRIVADO	270	14	M	MALA	11	14	09	10	10
NAKA6035	PRIVADO	270	14	F	MALA	14	16	15	13	14
PUJO6203	PRIVADO	270	14	M	MALA	18	20	19	18	17
REJO6040	PRIVADO	270	14	M	MALA	13	16	13	13	13
RIYHA6138	PRIVADO	270	14	F	MALA	16	19	19	18	16
RIMA6070	PRIVADO	270	14	F	MALA	13	17	13	13	12
SAIK6063	PRIVADO	270	14	M	MALA	17	19	19	18	18
SISE2020	PRIVADO	270	14	M	MALA	15	17	17	14	16
SOLU7135	PRIVADO	270	14	F	MALA	12	15	13	13	13
SOMI6151	PRIVADO	270	14	F	MALA	13	16	14	13	14
VEPA6138	PRIVADO	270	14	M	MALA	18	20	20	19	17

ALUMNO	Tipo de cole	Escala de pe	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
ASPI6208	Privado	S/ 360.00	12	M	MALA	17	14	18	16	15
CARCHA6261	Privado	S/ 360.00	12	M	MALA	11	11	12	11	9
CAROLI6328	Privado	S/ 360.00	12	M	MALA	16	16	18	15	16
CHUJI6317	Privado	S/ 360.00	12	F	MALA	15	13	15	11	12
ESAI6309	Privado	S/ 360.00	12	F	MALA	15	15	16	14	15
GAFA6328	Privado	S/ 360.00	12	M	MALA	15	14	17	15	15
JUDA6338	Privado	S/ 360.00	12	M	MALA	12	10	14	12	9
MALE6309	Privado	S/ 360.00	12	M	MALA	13	15	16	14	14
MALUZ7759	Privado	S/ 360.00	12	M	MALA	15	16	17	15	16
MOARI6309	Privado	S/ 360.00	12	F	MALA	13	15	17	14	14
PEWA7752	Privado	S/ 360.00	12	M	MALA	15	15	16	15	14
PILU8160	Privado	S/ 360.00	12	F	MALA	10	11	12	11	10
QUIRI6311	Privado	S/ 360.00	12	M	MALA	9	11	12	10	10
RAAD6370	Privado	S/ 360.00	12	M	MALA	16	16	18	16	17
RUSE6296	Privado	S/ 360.00	12	M	MALA	17	16	18	16	16
SALNA6309	Privado	S/ 360.00	12	F	MALA	13	11	14	12	9
SOLJU6311	Privado	S/ 360.00	12	M	MALA	11	11	11	10	10
TOCE6328	Privado	S/ 360.00	12	F	MALA	16	15	17	17	16
ZARHU7802	Privado	S/ 360.00	12	M	MALA	11	11	11	10	10
FRAKA6230	Privado	S/ 360.00	13	F	MALA	11	12	12	14	14
GODAN7479	Privado	S/ 360.00	13	F	MALA	14	15	14	16	16
HUANN6249	Privado	S/ 360.00	13	F	MALA	14	16	15	17	16
HUAJO6196	Privado	S/ 360.00	13	M	MALA	9	11	10	10	9
LANE6273	Privado	S/ 360.00	13	M	MALA	11	10	11	11	10
RAYO7564	Privado	S/ 360.00	13	M	MALA	14	13	13	14	13
ROED7530	Privado	S/ 360.00	13	M	MALA	12	12	12	13	13
RUDY7529	Privado	S/ 360.00	13	M	MALA	9	10	10	11	9
SALU6224	Privado	S/ 360.00	13	M	MALA	13	14	14	13	14
TOGRE6249	Privado	S/ 360.00	13	F	MALA	11	11	13	14	13
ACAL6238	Privado	S/ 360.00	14	M	MALA	13	12	11	12	12
ASZA6199	Privado	S/ 360.00	14	F	MALA	14	14	12	13	14
BALVI6038	Privado	S/ 360.00	14	M	MALA	19	18	19	18	19
CARES6196	Privado	S/ 360.00	14	F	MALA	15	16	16	15	14
DIJA6235	Privado	S/ 360.00	14	M	MALA	11	12	11	12	12
FELE6166	Privado	S/ 360.00	14	F	MALA	14	15	15	15	15
GAGI6198	Privado	S/ 360.00	14	M	MALA	14	13	13	13	13
LOZHA6198	Privado	S/ 360.00	14	F	MALA	17	16	17	16	17
MEDI6235	Privado	S/ 360.00	14	M	MALA	19	19	19	19	18
MOSE6198	Privado	S/ 360.00	14	M	MALA	11	12	12	10	11
NADA6207	Privado	S/ 360.00	14	M	MALA	11	11	11	10	11
PEAR2020	Privado	S/ 360.00	14	F	MALA	17	16	17	16	14
RONA6199	Privado	S/ 360.00	14	F	MALA	13	14	16	13	13
RUVI7384	Privado	S/ 360.00	14	M	MALA	9	8	9	10	8
TRUJE7355	Privado	S/ 360.00	14	M	MALA	16	16	15	15	14
VIRO6235	Privado	S/ 360.00	14	F	MALA	14	14	14	14	13
BEJI6152	Privado	S/ 360.00	15	F	MALA	16	15	14	13	14
CHUKA6165	Privado	S/ 360.00	15	F	MALA	15	15	14	13	14
CHUBRI6070	Privado	S/ 360.00	15	F	MALA	13	12	10	12	12
CODA6066	Privado	S/ 360.00	15	M	MALA	18	17	18	18	18
DINA6070	Privado	S/ 360.00	15	F	MALA	18	18	17	17	18
LALU6066	Privado	S/ 360.00	15	F	MALA	13	14	10	11	10
MABE7298	Privado	S/ 360.00	15	F	MALA	15	15	16	14	15
MAAN6070	Privado	S/ 360.00	15	M	MALA	9	11	9	10	9
MARPE6066	Privado	S/ 360.00	15	M	MALA	13	13	12	12	11
PUJO6203	Privado	S/ 360.00	15	M	MALA	18	18	17	19	18
REJO6040	Privado	S/ 360.00	15	M	MALA	11	13	13	12	13
RIMA6070	Privado	S/ 360.00	15	F	MALA	15	13	12	15	14
SAIK6063	Privado	S/ 360.00	15	M	MALA	16	17	17	17	18
VEPA6138	Privado	S/ 360.00	15	M	MALA	19	20	18	18	19

REGISTRO DE ASISTENCIA

AÑO 2024

AÑO	PERIODO	CODIGO_ALU	PRESENTE	TARDANZA I	TARDANZA J	INASISTENCI	INASISTENCIA JUSTIFICA
08/05/2024	PERIODO UNO	ASPI6208	23				
08/05/2024	PERIODO UNO	CARCHA6261	23				
08/05/2024	PERIODO UNO	CAROLI6328	13	6	2	1	
08/05/2024	PERIODO UNO	CHUJI6317	22	1			
08/05/2024	PERIODO UNO	ESAI6309	23				
08/05/2024	PERIODO UNO	GAFA6328	23				
08/05/2024	PERIODO UNO	JUDA6338	21	2			
08/05/2024	PERIODO UNO	MALE6309	12	11			
08/05/2024	PERIODO UNO	MALUZ7759	19	4			
08/05/2024	PERIODO UNO	MOARI6309	23				
08/05/2024	PERIODO UNO	PEWA7752	19	2			
08/05/2024	PERIODO UNO	PILU8160	11	12			
08/05/2024	PERIODO UNO	QUIRI6311	5	13			
08/05/2024	PERIODO UNO	RAAD6370	20	3			
08/05/2024	PERIODO UNO	RUSE6296	23				
08/05/2024	PERIODO UNO	SALNA6309	17	6			
08/05/2024	PERIODO UNO	SOLJU6311	22	1			
08/05/2024	PERIODO UNO	TOCE6328	22	1			
08/05/2024	PERIODO UNO	ZARHU7802	8	8			
08/05/2024	PERIODO UNO	FRAKA6230	22				
08/05/2024	PERIODO UNO	GODAN7479	19	3			
08/05/2024	PERIODO UNO	HUANN6249	22				
08/05/2024	PERIODO UNO	HUAJO6196	21	1			
08/05/2024	PERIODO UNO	LANE6273	14	8			
08/05/2024	PERIODO UNO	RAYO7564	19	3			
08/05/2024	PERIODO UNO	ROED7530	18	3			
08/05/2024	PERIODO UNO	RUDY7529	13	9			
08/05/2024	PERIODO UNO	SALU6224	22				
08/05/2024	PERIODO UNO	TOGRE6249	20	2			
08/05/2024	PERIODO UNO	ACAL6238	12	10			
08/05/2024	PERIODO UNO	ASZA6199	19	3	3		
08/05/2024	PERIODO UNO	BALVI6038	22				
08/05/2024	PERIODO UNO	CARES6196	12	7	2	1	
08/05/2024	PERIODO UNO	DIJA6235	20				
08/05/2024	PERIODO UNO	FELE6166	17	4	1		
08/05/2024	PERIODO UNO	GAGI6198	18	2			
08/05/2024	PERIODO UNO	LOZHA6198	14	8			
08/05/2024	PERIODO UNO	MEDI6235	22				
08/05/2024	PERIODO UNO	MOSE6198	2	18	2		
08/05/2024	PERIODO UNO	NADA6207	19	2		1	
08/05/2024	PERIODO UNO	PEAR2020	8	12	1		
08/05/2024	PERIODO UNO	RONA6199	20	2	1		
08/05/2024	PERIODO UNO	RUVI7384	8	13			
08/05/2024	PERIODO UNO	TRUJE7355	21	1			
08/05/2024	PERIODO UNO	VIRO6235	15	7			
08/05/2024	PERIODO UNO	BEJI6152	22				
08/05/2024	PERIODO UNO	CHUKA6165	22				
08/05/2024	PERIODO UNO	CHUBRI6070	19				
08/05/2024	PERIODO UNO	CODA6066	20				
08/05/2024	PERIODO UNO	DINA6070	22				
08/05/2024	PERIODO UNO	LALU6066	22				
08/05/2024	PERIODO UNO	MABE7298	21				
08/05/2024	PERIODO UNO	MAAN6070	15				
08/05/2024	PERIODO UNO	MARPE6066	18				
08/05/2024	PERIODO UNO	PUJO6203	21				
08/05/2024	PERIODO UNO	REJO6040	13				
08/05/2024	PERIODO UNO	RIMA6070	22				
08/05/2024	PERIODO UNO	SAIK6063	22				
08/05/2024	PERIODO UNO	VEPA6138	22				

AÑO 2023

08/05/2023	PERIODO UNO	CUWA6273	21	1				
08/05/2023	PERIODO UNO	FRAKA6230	22	1				
08/05/2023	PERIODO UNO	GODAN7479	22					
08/05/2023	PERIODO UNO	HUANN6249	35					
08/05/2023	PERIODO UNO	HUAJO6196	32	3				
08/05/2023	PERIODO UNO	LANE6273	32	2		1		
08/05/2023	PERIODO UNO	SALU6224	23	10		2		
08/05/2023	PERIODO UNO	TOGRE6249	26	6		3		
08/05/2023	PERIODO UNO	VIYA6257	33	2				
08/05/2023	PERIODO UNO	ACAL6238	32	2				
08/05/2023	PERIODO UNO	ASZA6199	29	6				
08/05/2023	PERIODO UNO	BALVI6038	6					
08/05/2023	PERIODO UNO	BRANI6199	1	4	1			
08/05/2023	PERIODO UNO	CARES6196	5		1			
08/05/2023	PERIODO UNO	DIJA6235	4	1				
08/05/2023	PERIODO UNO	FELE6166	2	3				
08/05/2023	PERIODO UNO	FRALE6038	6					
08/05/2023	PERIODO UNO	GAGI6198	2	4				
08/05/2023	PERIODO UNO	GUAGU6238	5					
08/05/2023	PERIODO UNO	HERI6216	6					
08/05/2023	PERIODO UNO	HUASHAN74	5					
08/05/2023	PERIODO UNO	LOZHA6198	4	1				
08/05/2023	PERIODO UNO	MOSE6198	2	4				
08/05/2023	PERIODO UNO	NADA6207	6					
08/05/2023	PERIODO UNO	PEAR2020	4	1				
08/05/2023	PERIODO UNO	TEMO7384	6					
08/05/2023	PERIODO UNO	TRUJE7355	5					
08/05/2023	PERIODO UNO	VAYO6238	4	2				
08/05/2023	PERIODO UNO	VIRO6235	4	3				
08/05/2023	PERIODO UNO	ARJE6202	4	2				
08/05/2023	PERIODO UNO	CAAN6149	12	5			1	
08/05/2023	PERIODO UNO	CECI6066	29	2				
08/05/2023	PERIODO UNO	CHUKA6165	31					
08/05/2023	PERIODO UNO	CODA6066	23	1		1	1	
08/05/2023	PERIODO UNO	DINA6070	31					
08/05/2023	PERIODO UNO	LALU6066	30				1	
08/05/2023	PERIODO UNO	MABE7298	20	9		1	1	
08/05/2023	PERIODO UNO	MAAN6070	15	15		1		
08/05/2023	PERIODO UNO	NAKA6035	27	4				
08/05/2023	PERIODO UNO	PUJO6203	31					
08/05/2023	PERIODO UNO	REJO6040	28	2		1		
08/05/2023	PERIODO UNO	RIYHA6138	30	1				
08/05/2023	PERIODO UNO	RIMA6070	21	3			2	
08/05/2023	PERIODO UNO	SAIK6063	30	1				
08/05/2023	PERIODO UNO	SISE2020	17	7		1		
08/05/2023	PERIODO UNO	SOLU7135	31					
08/05/2023	PERIODO UNO	SOMI6151	22	6			3	
08/05/2023	PERIODO UNO	VEPA6138	31					

25/07/2023	PERIODO DOS	CUWA6273	31	8				
25/07/2023	PERIODO DOS	FRAKA6230	32	1				5
25/07/2023	PERIODO DOS	GODAN7479	38					
25/07/2023	PERIODO DOS	HUANN6249	35	1	2			
25/07/2023	PERIODO DOS	HUAJO6196	34	3		1		
25/07/2023	PERIODO DOS	LANE6273	32	6				
25/07/2023	PERIODO DOS	PABRI6217	17	14		6		
25/07/2023	PERIODO DOS	ROED7530	30	4				
25/07/2023	PERIODO DOS	SALU6224	33	4				1
25/07/2023	PERIODO DOS	TOGRE6249	37					1
25/07/2023	PERIODO DOS	VIYA6257	20	14	1	2		
25/07/2023	PERIODO DOS	ACAL6238	22	12				
25/07/2023	PERIODO DOS	ASZA6199	29	2	1			2
25/07/2023	PERIODO DOS	BALVI6038	29	1	1			3
25/07/2023	PERIODO DOS	BRANI6199	12	13		6		2
25/07/2023	PERIODO DOS	CARES6196	17	14	2			
25/07/2023	PERIODO DOS	DIJA6235	25	6	1	1		
25/07/2023	PERIODO DOS	FELE6166	13	17		3		
25/07/2023	PERIODO DOS	FRALE6038	33	1				
25/07/2023	PERIODO DOS	GAGI6198	22	12				
25/07/2023	PERIODO DOS	GUAGU6238	29	5				
25/07/2023	PERIODO DOS	HERI6216	32	1		1		
25/07/2023	PERIODO DOS	HUASHAN74	26	2		6		
25/07/2023	PERIODO DOS	LOZHA6198	18	14	2			
25/07/2023	PERIODO DOS	MOSE6198	13	18		1		2
25/07/2023	PERIODO DOS	NADA6207	21	10	2	1		
25/07/2023	PERIODO DOS	PEAR2020	15	13				5
25/07/2023	PERIODO DOS	TEMO7384	34					
25/07/2023	PERIODO DOS	TRUJE7355	30	1		2		
25/07/2023	PERIODO DOS	VAYO6238	19	11				3
25/07/2023	PERIODO DOS	VIRO6235	19	12	1	1		
25/07/2023	PERIODO DOS	ARJE6202	8	10		2		
25/07/2023	PERIODO DOS	CAAN6149	12	5		1		
25/07/2023	PERIODO DOS	CECI6066	16	3	1			
25/07/2023	PERIODO DOS	CHUKA6165	18	2				
25/07/2023	PERIODO DOS	CODA6066	16	4				
25/07/2023	PERIODO DOS	DINA6070	20					
25/07/2023	PERIODO DOS	LALU6066	20					
25/07/2023	PERIODO DOS	MABE7298	16	2		1		1
25/07/2023	PERIODO DOS	MAAN6070	12	5		1		2
25/07/2023	PERIODO DOS	NAKA6035	16	4				
25/07/2023	PERIODO DOS	PUJO6203	18			1		1
25/07/2023	PERIODO DOS	REJO6040	18			1		1
25/07/2023	PERIODO DOS	RIYHA6138	18					2
25/07/2023	PERIODO DOS	RIMA6070	11	4	1	1		3
25/07/2023	PERIODO DOS	SAIK6063	20					
25/07/2023	PERIODO DOS	SISE2020	15	5				
25/07/2023	PERIODO DOS	SOLU7135	18			1		1
25/07/2023	PERIODO DOS	SOMI6151	10	7	2	1		
25/07/2023	PERIODO DOS	VEPA6138	20					

06/10/2023	PERIODO TRES	CUWA6273	33	8		2	1
06/10/2023	PERIODO TRES	FRAKA6230	42	1			
06/10/2023	PERIODO TRES	GODAN7479	40	3			
06/10/2023	PERIODO TRES	HUANN6249	40	3			
06/10/2023	PERIODO TRES	HUAJO6196	28	9		6	
06/10/2023	PERIODO TRES	LANE6273	43				
06/10/2023	PERIODO TRES	PABRI6217	31	11		1	
06/10/2023	PERIODO TRES	ROED7530	39	3		1	
06/10/2023	PERIODO TRES	SALU6224	38	5			
06/10/2023	PERIODO TRES	TOGRE6249	41	2			
06/10/2023	PERIODO TRES	VIYA6257	17	25		1	
06/10/2023	PERIODO TRES	ACAL6238	24	3			1
06/10/2023	PERIODO TRES	ASZA6199	23	3	1	1	
06/10/2023	PERIODO TRES	BALVI6038	22		2		4
06/10/2023	PERIODO TRES	BRANI6199	9	15	1	3	
06/10/2023	PERIODO TRES	CARES6196	3	23			2
06/10/2023	PERIODO TRES	DIJA6235	26	2			
06/10/2023	PERIODO TRES	FELE6166	14	14			
06/10/2023	PERIODO TRES	FRALE6038	23	5			
06/10/2023	PERIODO TRES	GAGI6198	21	7			
06/10/2023	PERIODO TRES	GUAGU6238	22	6			
06/10/2023	PERIODO TRES	HERI6216	25	2			1
06/10/2023	PERIODO TRES	LOZHA6198	7	19	2		
06/10/2023	PERIODO TRES	MOSE6198	9	18			1
06/10/2023	PERIODO TRES	NADA6207	19	7			2
06/10/2023	PERIODO TRES	PEAR2020	11	17			
06/10/2023	PERIODO TRES	TEMO7384	25	3			
06/10/2023	PERIODO TRES	TRUJE7355	28				
06/10/2023	PERIODO TRES	VAYO6238	21	7			
06/10/2023	PERIODO TRES	VIRO6235	21	6		1	
06/10/2023	PERIODO TRES	ARJE6202	36	2		2	
06/10/2023	PERIODO TRES	CAAN6149	33	1	1	2	2
06/10/2023	PERIODO TRES	CECI6066	24	6		4	5
06/10/2023	PERIODO TRES	CHUKA6165	40				
06/10/2023	PERIODO TRES	CHUBRI6070	25	1		1	
06/10/2023	PERIODO TRES	CODA6066	38	2			
06/10/2023	PERIODO TRES	DINA6070	39				1
06/10/2023	PERIODO TRES	LALU6066	40				
06/10/2023	PERIODO TRES	MABE7298	31	3			4
06/10/2023	PERIODO TRES	MAAN6070	25	12		2	1
06/10/2023	PERIODO TRES	NAKA6035	35	4			1
06/10/2023	PERIODO TRES	PUJO6203	39				1
06/10/2023	PERIODO TRES	REJO6040	32	2	3	3	
06/10/2023	PERIODO TRES	RIYHA6138	40				
06/10/2023	PERIODO TRES	RIMA6070	32	6		2	
06/10/2023	PERIODO TRES	SAIK6063	39				1
06/10/2023	PERIODO TRES	SISE2020	32	7		1	
06/10/2023	PERIODO TRES	SOLU7135	38	1			
06/10/2023	PERIODO TRES	SOMI6151	30	5	2	1	2
06/10/2023	PERIODO TRES	VEPA6138	40				

21/12/2023	PERIODO CUA	CUWA6273	21	3				
21/12/2023	PERIODO CUA	FRAKA6230	23				1	
21/12/2023	PERIODO CUA	GODAN7479	20	4				
21/12/2023	PERIODO CUA	HUANN6249	24					
21/12/2023	PERIODO CUA	HUAJO6196	13	3		8		
21/12/2023	PERIODO CUA	LANE6273	23			1		
21/12/2023	PERIODO CUA	PABRI6217	21	1		2		
21/12/2023	PERIODO CUA	ROED7530	17	5		2		
21/12/2023	PERIODO CUA	SALU6224	19	5				
21/12/2023	PERIODO CUA	TEDA6250	5	13		6		
21/12/2023	PERIODO CUA	VIYA6257	11	10		3		
21/12/2023	PERIODO CUA	ACAL6238	30	1			1	
21/12/2023	PERIODO CUA	ASZA6199	13	6	1	1	1	
21/12/2023	PERIODO CUA	BALVI6038	21		1			
21/12/2023	PERIODO CUA	BRANI6199	6	9		5	1	
21/12/2023	PERIODO CUA	CARES6196	5	15		1		
21/12/2023	PERIODO CUA	DIJA6235	22					
21/12/2023	PERIODO CUA	FELE6166	10	10	1		1	
21/12/2023	PERIODO CUA	FRALE6038	18	3		1		
21/12/2023	PERIODO CUA	GAGI6198	15	5			1	
21/12/2023	PERIODO CUA	GUAGU6238	13	8				
21/12/2023	PERIODO CUA	HERI6216	18	4				
21/12/2023	PERIODO CUA	LOZHA6198	6	11	1	3		
21/12/2023	PERIODO CUA	MOSE6198	5	15		1		
21/12/2023	PERIODO CUA	NADA6207	14	5	1	1		
21/12/2023	PERIODO CUA	PEAR2020	8	12		1		
21/12/2023	PERIODO CUA	TEMO7384	22					
21/12/2023	PERIODO CUA	TRUJE7355	18	4				
21/12/2023	PERIODO CUA	VAYO6238	16	5		1		
21/12/2023	PERIODO CUA	VIRO6235	11	9		2		
21/12/2023	PERIODO CUA	ARJE6202	16	2	1	2	2	
21/12/2023	PERIODO CUA	CAAN6149	12	5		1	5	
21/12/2023	PERIODO CUA	CECI6066	13	5		3	1	
21/12/2023	PERIODO CUA	CHUKA6165	22	1				
21/12/2023	PERIODO CUA	CHUBRI6070	17	4		2		
21/12/2023	PERIODO CUA	CODA6066	16	7				
21/12/2023	PERIODO CUA	DINA6070	23					
21/12/2023	PERIODO CUA	LALU6066	22	1				
21/12/2023	PERIODO CUA	MABE7298	15	5		2	1	
21/12/2023	PERIODO CUA	MAAN6070	5	16		1		
21/12/2023	PERIODO CUA	NAKA6035	15	8				
21/12/2023	PERIODO CUA	PUJO6203	23					
21/12/2023	PERIODO CUA	REJO6040	19	2		2		
21/12/2023	PERIODO CUA	RIYHA6138	21	2				
21/12/2023	PERIODO CUA	RIMA6070	13	9		1		
21/12/2023	PERIODO CUA	SAIK6063	22	1				
21/12/2023	PERIODO CUA	SISE2020	19	4				
21/12/2023	PERIODO CUA	SOLU7135	21			1	1	
21/12/2023	PERIODO CUA	SOMI6151	17	5	1			
21/12/2023	PERIODO CUA	VEPA6138	23					

AÑO 2022

13/05/2022	PERIODO UNO	ACAL6238	19				
13/05/2022	PERIODO UNO	ASZA6199	16	2			1
13/05/2022	PERIODO UNO	BALVI6038	16	2	1		
13/05/2022	PERIODO UNO	CARES6196	1	8	9		
13/05/2022	PERIODO UNO	CHUASH2020	11	8			
13/05/2022	PERIODO UNO	DIJA6235	17	2			
13/05/2022	PERIODO UNO	FELE6166	12	5	1		
13/05/2022	PERIODO UNO	FRALE6038	18				1
13/05/2022	PERIODO UNO	HUASHAN74	13	5			1
13/05/2022	PERIODO UNO	LOZHA6198	16	1			1
13/05/2022	PERIODO UNO	MOSE6198	16	2			1
13/05/2022	PERIODO UNO	TEJU7384	19				
13/05/2022	PERIODO UNO	TRUJE7355	19				
13/05/2022	PERIODO UNO	ARJE6202	25	3			
13/05/2022	PERIODO UNO	CALAN6149	19				
13/05/2022	PERIODO UNO	CHUKA6165	26				
13/05/2022	PERIODO UNO	CODA6066	26	1			
13/05/2022	PERIODO UNO	DINA6070	28				
13/05/2022	PERIODO UNO	LALU6066	27		1		
13/05/2022	PERIODO UNO	MAAN6070	12	15			1
13/05/2022	PERIODO UNO	NAKA6035	25	1			
13/05/2022	PERIODO UNO	PUJO6203	26			1	
13/05/2022	PERIODO UNO	REJO6040	23	1	3		
13/05/2022	PERIODO UNO	RIYHA6138	17				
13/05/2022	PERIODO UNO	RIMA6070	22	6			
13/05/2022	PERIODO UNO	SAIK6063	28				
13/05/2022	PERIODO UNO	SISE2020	19	6			1
13/05/2022	PERIODO UNO	SOMI6151	27		1		
13/05/2022	PERIODO UNO	VEPA6138	28				

26/07/2022	PERIODO DOS	ACAL6238	37	6	1		1
26/07/2022	PERIODO DOS	ASZA6199	40		1		4
26/07/2022	PERIODO DOS	BALVI6038	39				6
26/07/2022	PERIODO DOS	CARES6196	8	15	21		1
26/07/2022	PERIODO DOS	CHUASH2020	35	7	1		2
26/07/2022	PERIODO DOS	DIJA6235	44				1
26/07/2022	PERIODO DOS	FELE6166	30	13	2		
26/07/2022	PERIODO DOS	FRALE6038	45				
26/07/2022	PERIODO DOS	GUAG6238	22				2
26/07/2022	PERIODO DOS	HUASHAN74	31	12			2
26/07/2022	PERIODO DOS	LOZHA6198	41	4			
26/07/2022	PERIODO DOS	MOSE6198	40	1		1	3
26/07/2022	PERIODO DOS	TEJU7384	42				3
26/07/2022	PERIODO DOS	TRUJE7355	43				2
26/07/2022	PERIODO DOS	ARJE6202	24	11	2		5
26/07/2022	PERIODO DOS	CALAN6149	39	1		1	1
26/07/2022	PERIODO DOS	CHUKA6165	42				
26/07/2022	PERIODO DOS	CODA6066	34	8			
26/07/2022	PERIODO DOS	DINA6070	40	1			1
26/07/2022	PERIODO DOS	LALU6066	40	2			
26/07/2022	PERIODO DOS	MAAN6070	17	21	1		2
26/07/2022	PERIODO DOS	NAKA6035	40		2		
26/07/2022	PERIODO DOS	PUJO6203	41	1			
26/07/2022	PERIODO DOS	REJO6040	27	9	5		
26/07/2022	PERIODO DOS	RIYHA6138	40		2		
26/07/2022	PERIODO DOS	RIMA6070	25	12	4		
26/07/2022	PERIODO DOS	SAIK6063	42				
26/07/2022	PERIODO DOS	SISE2020	35	7			
26/07/2022	PERIODO DOS	SOMI6151	39	2			
26/07/2022	PERIODO DOS	VEPA6138	42				

07/10/2022	PERIODO TRES	ACAL6238	28	10			
07/10/2022	PERIODO TRES	ASZA6199	36	1	1		
07/10/2022	PERIODO TRES	BALVI6038	38				
07/10/2022	PERIODO TRES	BRANI6199	25	7	3		
07/10/2022	PERIODO TRES	CARES6196	1	1	36		
07/10/2022	PERIODO TRES	CHUASH2020	27	11			
07/10/2022	PERIODO TRES	DIJA6235	37			1	
07/10/2022	PERIODO TRES	FELE6166	27	9	1		1
07/10/2022	PERIODO TRES	FRALE6038	36	2			
07/10/2022	PERIODO TRES	GUAG6238	36	2			
07/10/2022	PERIODO TRES	HUASHAN74	19	17	2		
07/10/2022	PERIODO TRES	LOZHA6198	31	6	1		
07/10/2022	PERIODO TRES	MOSE6198	37	1			
07/10/2022	PERIODO TRES	TEJU7384	38				
07/10/2022	PERIODO TRES	TRUJE7355	38				
07/10/2022	PERIODO TRES	ARJE6202	13	13		4	1
07/10/2022	PERIODO TRES	CALAN6149	25	1	1	4	
07/10/2022	PERIODO TRES	CHUKA6165	31				
07/10/2022	PERIODO TRES	CODA6066	22	9			
07/10/2022	PERIODO TRES	DINA6070	31				
07/10/2022	PERIODO TRES	LALU6066	30				1
07/10/2022	PERIODO TRES	MAAN6070	25	3		2	1
07/10/2022	PERIODO TRES	NAKA6035	31	1			
07/10/2022	PERIODO TRES	PUJO6203	31				
07/10/2022	PERIODO TRES	REJO6040	18	11			2
07/10/2022	PERIODO TRES	RIYHA6138	31				
07/10/2022	PERIODO TRES	RIMA6070	13	16	2		
07/10/2022	PERIODO TRES	SAIK6063	31				
07/10/2022	PERIODO TRES	SISE2020	26	5			
07/10/2022	PERIODO TRES	SOMI6151	18	12			1
07/10/2022	PERIODO TRES	VEPA6138	31				

20/12/2022	PERIODO CUAT	ACAL6238	28	7	1		
20/12/2022	PERIODO CUAT	ASZA6199	32	3			1
20/12/2022	PERIODO CUAT	BALVI6038	36				
20/12/2022	PERIODO CUAT	BRANI6199	10	22	3	1	
20/12/2022	PERIODO CUAT	CARES6196	4	30	2		
20/12/2022	PERIODO CUAT	CHUASH2020	25	11			
20/12/2022	PERIODO CUAT	DIJA6235	34	1	1		
20/12/2022	PERIODO CUAT	FELE6166	24	12			
20/12/2022	PERIODO CUAT	FRALE6038	35	1			
20/12/2022	PERIODO CUAT	GUAG6238	36				
20/12/2022	PERIODO CUAT	HUASHAN74	31	5			
20/12/2022	PERIODO CUAT	LOZHA6198	28	8			
20/12/2022	PERIODO CUAT	MOSE6198	34		1	1	
20/12/2022	PERIODO CUAT	TEJU7384	35				1
20/12/2022	PERIODO CUAT	TRUJE7355	33	3			
20/12/2022	PERIODO CUAT	ARJE6202	5	20		1	5
20/12/2022	PERIODO CUAT	CALAN6149	22	7			2
20/12/2022	PERIODO CUAT	CHUKA6165	29	2			
20/12/2022	PERIODO CUAT	CODA6066	29	2			
20/12/2022	PERIODO CUAT	DINA6070	29	2			
20/12/2022	PERIODO CUAT	LALU6066	31				
20/12/2022	PERIODO CUAT	MAAN6070	25	5			1
20/12/2022	PERIODO CUAT	NAKA6035	23	7			1
20/12/2022	PERIODO CUAT	PUJO6203	29	1			1
20/12/2022	PERIODO CUAT	REJO6040	21	9	1		
20/12/2022	PERIODO CUAT	RIYHA6138	30	1			
20/12/2022	PERIODO CUAT	SAIK6063	29	2			
20/12/2022	PERIODO CUAT	SISE2020	16	13		1	1
20/12/2022	PERIODO CUAT	SOMI6151	18	11	1		1
20/12/2022	PERIODO CUAT	VEPA6138	31				

AÑO 2021

25/05/2021	PERIODO UNO	ARJE6202	10	14		2	
25/05/2021	PERIODO UNO	CALAN6149	21	2	1		2
25/05/2021	PERIODO UNO	CODA6066	25		1		
25/05/2021	PERIODO UNO	DINA6070	25				1
25/05/2021	PERIODO UNO	MAAN6070	8	17			1
25/05/2021	PERIODO UNO	PUJO6203	25	1			
25/05/2021	PERIODO UNO	REJO6040	24	2			
25/05/2021	PERIODO UNO	RIYHA6138	20	6			
25/05/2021	PERIODO UNO	RIMA6070	10	14	2		
25/05/2021	PERIODO UNO	SAIK6063	20	5	1		
25/05/2021	PERIODO UNO	SOMI6151	25	1			
25/05/2021	PERIODO UNO	VAJO6203	16	5	2		1
25/05/2021	PERIODO UNO	VEPA6138	26				
25/05/2021	PERIODO UNO	VILA6007	7	19			
10/09/2021	PERIODO DOS	ARJE6202	30	14	1		
10/09/2021	PERIODO DOS	CALAN6149	30	10	3	1	1
10/09/2021	PERIODO DOS	CODA6066	44	1			
10/09/2021	PERIODO DOS	DINA6070	45				
10/09/2021	PERIODO DOS	LALU6066	38	1	2		
10/09/2021	PERIODO DOS	MAAN6070	23	22			
10/09/2021	PERIODO DOS	PUJO6203	35	6	3		1
10/09/2021	PERIODO DOS	REJO6040	38	6	1		
10/09/2021	PERIODO DOS	RIYHA6138	41	2	1		1
10/09/2021	PERIODO DOS	RIMA6070	17	25	3		
10/09/2021	PERIODO DOS	SAIK6063	36	5	4		
10/09/2021	PERIODO DOS	SOMI6151	35	4	2	2	2
10/09/2021	PERIODO DOS	VAJO6203	27	16	2		
10/09/2021	PERIODO DOS	VEPA6138	43	1			1
10/09/2021	PERIODO DOS	VILA6007	17	19		2	7
17/12/2021	PERIODO TRES	ARJE6202	19	33			
17/12/2021	PERIODO TRES	CALAN6149	40	12			
17/12/2021	PERIODO TRES	CODA6066	52				
17/12/2021	PERIODO TRES	DINA6070	50	2			
17/12/2021	PERIODO TRES	LALU6066	51		1		
17/12/2021	PERIODO TRES	MAAN6070	38	13	1		
17/12/2021	PERIODO TRES	PUJO6203	49	3			
17/12/2021	PERIODO TRES	REJO6040	33	16	2		1
17/12/2021	PERIODO TRES	RIYHA6138	45	5	2		
17/12/2021	PERIODO TRES	RIMA6070	29	23			
17/12/2021	PERIODO TRES	SAIK6063	44	8			
17/12/2021	PERIODO TRES	SOMI6151	40	9	2		1
17/12/2021	PERIODO TRES	VAJO6203	41	10			1
17/12/2021	PERIODO TRES	VEPA6138	52				
17/12/2021	PERIODO TRES	VILA6007	18	32		2	

Anexo 6: Autorización de la Institución para realizar la investigación

Autorización para publicar los datos de la entidad en el repositorio

De conformidad con la Ley de derechos de autor Decreto legislativo N° 822, la Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto Ley N° 30035 y la Directiva que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto (ALICIA) administrado por el Pliego Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica – CONCYTEC con Resolución De Presidencia N° 048-2020-CONCYTEC-P;

Autorizo	(X)	No Autorizo	( )

Publicar la identidad de la organización, en la cual se lleva a cabo la investigación y autorizo se haga el depósito del trabajo de investigación en el repositorio institucional de la Universidad Nacional de Cañete.

Nombre del trabajo de Investigación	
Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete	
Nombre de la Escuela Profesional	
Escuela Profesional de Ingeniería de Sistemas	
Autor	**DNI**
Bach. Chumpitaz Luna Jeanice Daniela	78549112
Bach. Hinojosa Rojas Anel Cecilia	76302907

Entiendo que la investigación se guardará en el Repositorio Institucional de la Universidad Nacional de Cañete, estará disponible para los usuarios y se podrá citar en futuras investigaciones, con el reconocimiento de los derechos de propiedad intelectual del autor (es) del estudio.

Nombre de la Organización	**RUC**	**Firma y Sello**
I. E. MIS PRIMEROS PASOS		
Nombre del Titular o representante Legal	**DNI**	
Rosa Lezmila Evangelista Trabejo	15035203	

San Vicente de Cañete, 09 de julio del 2024

Anexo 7: Metodología de desarrollo de la solución

La elección de la metodología CRISP-DM como fundamento para el desarrollo de esta investigación se basó en la evaluación comparativa de dos principales metodologías mediante una tabla, como se observa en la Tabla 2.

Tabla 2: Cuadro comparativo de metodologías para minería de datos

	CRISP-DM	**KDD**
Enfoque	Específico para proyectos de minería de datos	Enfoque más amplio que cubre todo el proceso de descubrimiento de conocimiento en bases de datos
Fases Principales	(1) Comprensión del negocio (2) Comprensión de los datos (3) Preparación de los d (4) Modelado (5) Evaluación (6) Despliegue	(1) Selección de datos (2) Preprocesamiento (3) Transformación (4) Minería de datos (5) Interpretación/Evaluación de patrones (6) Utilización de conocimiento descubierto
Flexibilidad	Estructurado y adecuadodetallado, para proyectos específicos	Más flexible y adaptable, puede ser extendido según necesidades específicas y contextos
Uso	Ampliamente utilizado en la industria para proyectos de minería de datos	Más comúnmente utilizado en la investigación y desarrollo en bases de datos y minería de datos
Popularidad	Muy popular en aplicaciones comerciales y académicas	Menos estructurado pero influyente en la investigación y desarrollo de técnicas avanzadas
Etapas detalladas	Proporciona pasos específicos y detallados para cada fase del proceso de minería de datos	Más general en la descripción de las etapas, permitiendo adaptación según necesidades específicas
Aplicación Práctica	Enfoque práctico para la implementación eficiente de proyectos de minería de datos	Enfoque más teórico y exploratorio, adecuado para la investigación en descubrimiento de conocimiento
Foco Principal	Orientado a la implementación efectiva y eficiente de proyectos de minería de datos	Orientado a la investigación y desarrollo de técnicas avanzadas para descubrimiento de conocimiento

Ambas metodologías tienen sus puntos fuertes y están diseñadas para satisfacer distintas necesidades en el proceso de análisis y descubrimiento de conocimiento a partir de datos. Por las razones expuestas, se optó por utilizar la metodología CRISP-DM debido a que proporciona un marco estructurado adaptable para proyectos específicos. Además, ofrece una estructura clara y

bien definida para cada fase del proceso de minería de datos, lo que simplifica la gestión y ejecución del proyecto. Esta metodología también es lo suficientemente flexible como para ajustarse a diferentes contextos y tipos de datos, incluidos aquellos en el ámbito educativo, que suelen ser complejos y diversos.

METODOLOGÍA CRISP-DM

ETAPA 1: COMPRENSIÓN DEL NEGOCIO

En esta primera etapa se recolectó la información de los estudiantes del nivel Secundaria de la Institución Educativa Mis Primeros Pasos-Cañete, conociendo sus objetivos y analizando los factores que influyen en el rendimiento académico de los estudiantes.

a) Determinación del objetivo de negocio.

La Institución Educativa Mis Primeros Pasos del distrito de Mala de la provincia de Cañete es una comunidad integrada por alumnos, docentes, personal administrativo que brinda servicios educativos, lo cual consta de tres niveles: inicial, primaria y secundaria. Brinda una educación integral de nuestros alumnos mediante la formación integral de la más alta calidad al servicio de la niñez y la juventud, con exigencia académica bajo un clima de convivencia escolar democrático, fomentando prácticas y relaciones solidarias; Entre sus objetivos está mejorar la calidad académica de los alumnos y egresados.

Figura 35: Insignia de la Institución Educativa

b) Evaluación de la situación.

La Institución Educativa, que apuesta por brindar servicios educativos, ha enfrentado algunos desafíos en materia de calidad de la educación, pues el rendimiento académico de los estudiantes el año pasado fue insuficiente en comparación con el momento de la pandemia de coronavirus de 2019.

Dicha Institución carece de una información detallada, pero sí dispersa manualmente ya sea: archivos digitales, seguimiento escolar en los docentes de la Institución. Este

problema en los estudiantes podría afectar en su conducta o entre otros factores y sobre todo la pérdida del año escolar.

Los datos de los alumnos desde el momento de la matricula del hijo (a) a la Institución, los cuales contienen datos como Apellidos y nombres, edad, sexo, tipo de colegio, lugar de procedencia. Las notas de los alumnos se encuentran en la plataforma de los docentes. Todos los atributos han sido almacenaos en Excel y se encuentran centralizado en oficina de Dirección.

Se dispone de los datos de los alumnos ingresante desde el 2021 hasta 2024 que hacen un total de 2110 registros de notas con sus atributos.

c) Determinación de los objetivos de minería de datos.

El objetivo de la minería de datos es dar soporte mediante las técnicas de minería de datos a los objetivos de la investigación.

Objetivo general:

Determinar la predicción del rendimiento académico de los estudiantes de la institución educativa Mis Primeros Pasos, Cañete. Objetivos específicos:

- Determinar los indicadores de impacto que ayuda a reconocer los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete.
- Determinar el algoritmo óptimo que genere un modelo para predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.
- Determinar una predicción al rendimiento académico que ayudará con el aprendizaje personalizado a los estudiantes de la Institución Educativa Mis Primeros Pasos.

ETAPA 2: COMPRENSIÓN DE LOS DATOS

En esta segunda fase, toda la información de datos brindada de la Institución Educativa las cuales corresponden a los estudiantes del Nivel secundario se trabajó con los registros como asistencias y de notas) que contiene de cada estudiante del curso de Matemáticas durante los años 2021 para su posterior análisis. Como se puede visualizar en el Anexo 5.

a) Recolección de datos iniciales:

Estos datos se recolectaron en hoja de cálculo Excel conteniendo los siguientes campos como se visualiza en la Figura 36:

APELLIDOS Y NOMBRES DEL ALUMNO

TIPO DE COLEGIO

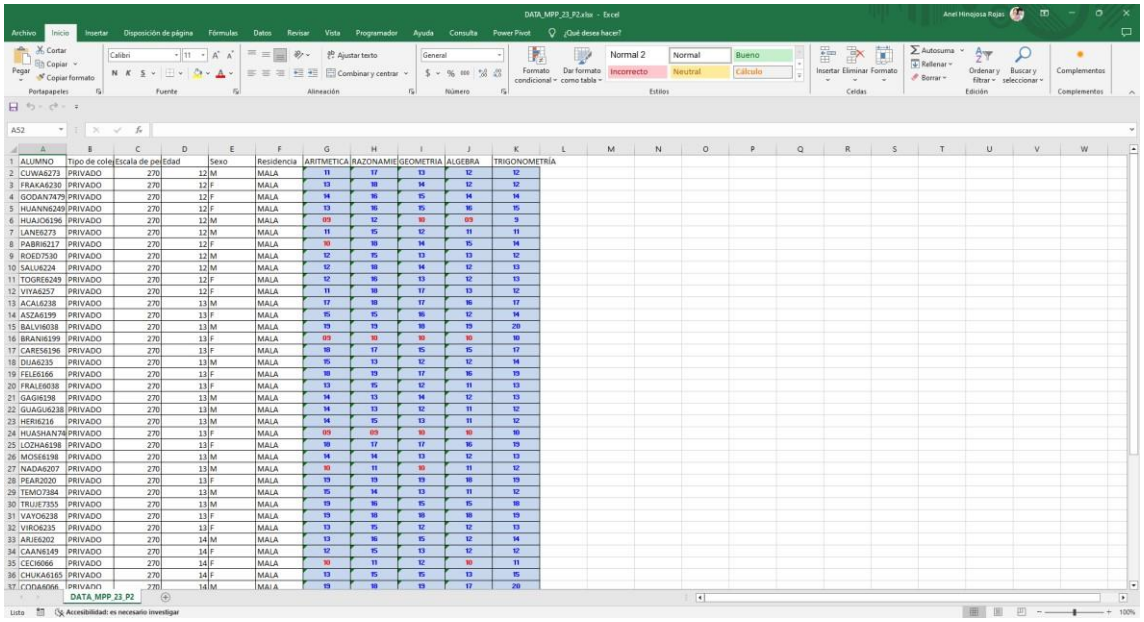
ESCALA DE PENSIÓN

EDAD

SEXO

RESIDENCIA

NOTAS DE MATEMÁTICA (Aritmética, Razonamiento Matemático, Geometría, Algebra y Trigonometría)



ALUMNO	Tipo de cole	Escala de per	Edad	Sexo	Residencia	ARITMETICA	RAZONAMIE	GEOMETRIA	ALGEBRA	TRIGONOMETRÍA
CUWA6273	PRIVADO	270	12	M	MALA	11	17	13	12	12
FRAKA6230	PRIVADO	270	12	F	MALA	13	18	14	12	12
GODAN7479	PRIVADO	270	12	F	MALA	14	16	15	14	14
HUANN6249	PRIVADO	270	12	F	MALA	13	16	15	16	15
HUAJO6196	PRIVADO	270	12	M	MALA	03	12	10	03	3
LANE6273	PRIVADO	270	12	M	MALA	11	15	12	11	11
PABRI6217	PRIVADO	270	12	F	MALA	10	18	14	15	14
ROED7530	PRIVADO	270	12	M	MALA	12	15	13	13	12
SALU6224	PRIVADO	270	12	M	MALA	12	18	14	12	13
TOGRE6249	PRIVADO	270	12	F	MALA	12	16	13	12	13
VIYA6257	PRIVADO	270	12	F	MALA	11	18	17	13	12
ACAL6238	PRIVADO	270	13	M	MALA	17	18	17	16	17
ASZA6199	PRIVADO	270	13	F	MALA	15	15	16	12	14
BALVI6038	PRIVADO	270	13	M	MALA	19	19	18	19	20
BRANI6199	PRIVADO	270	13	F	MALA	09	10	10	10	10
CARES6196	PRIVADO	270	13	F	MALA	18	17	15	15	17
DIJA6235	PRIVADO	270	13	M	MALA	15	13	12	12	14
FELE6166	PRIVADO	270	13	F	MALA	18	19	17	16	13
FRAL6038	PRIVADO	270	13	F	MALA	13	15	12	11	13
GAGI6198	PRIVADO	270	13	M	MALA	14	13	14	12	13
GUAGU6238	PRIVADO	270	13	M	MALA	14	13	12	11	12
HERI6216	PRIVADO	270	13	M	MALA	14	15	13	11	12
HUASHAN74	PRIVADO	270	13	F	MALA	09	09	10	10	10
LOZHA6198	PRIVADO	270	13	F	MALA	18	17	17	16	15
MOSI6198	PRIVADO	270	13	M	MALA	14	14	13	12	13
NADA6207	PRIVADO	270	13	M	MALA	10	11	10	11	12
PEAR2020	PRIVADO	270	13	F	MALA	19	19	19	18	19
TEMO7384	PRIVADO	270	13	M	MALA	15	14	13	11	12
TRUJE7355	PRIVADO	270	13	M	MALA	19	16	15	15	18
VAYO6238	PRIVADO	270	13	F	MALA	19	18	18	18	19
VIRO6235	PRIVADO	270	13	F	MALA	13	15	12	12	13
ARJE6202	PRIVADO	270	14	M	MALA	13	16	15	12	14
CAAN6149	PRIVADO	270	14	F	MALA	12	15	13	12	12
CECI6066	PRIVADO	270	14	F	MALA	10	11	12	10	11
CHUKA6165	PRIVADO	270	14	F	MALA	13	15	15	13	15
CODA6066	PRIVADO	270	14	M	MALA	19	18	19	17	20

Figura 36: Registro de notas en Excel

Los datos de ASISTENCIAS también se recolectaron en un Excel con los siguientes atributos, como se visualiza en la Figura 37:

AÑO PERIODO

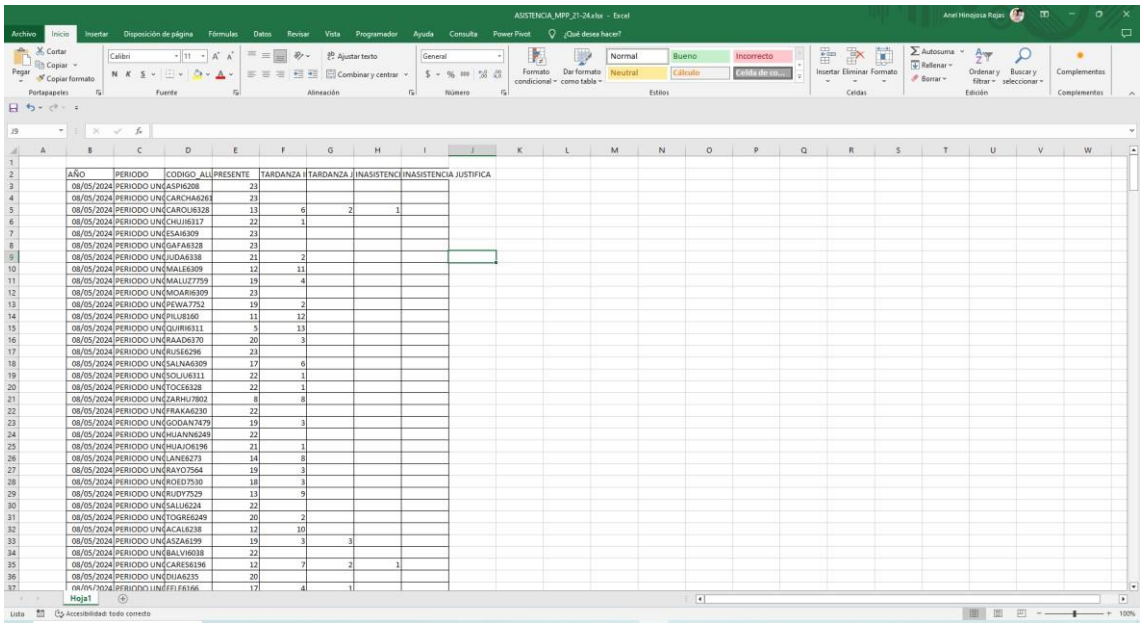
CODIGO_ALUMNO

PRESENTE

TARDANZA INJUSTIFICADA

TARDANZA JUSTIFICADA

INASISTENCIA INJUSTIFICADA
INASISTENCIA JUSTIFICADA



AÑO	PERIODO	CODIGO_ALU	PRESENTE	TARDANZA J	TARDANZA J	INASISTENCI	INASISTENCIA JUSTIFICA
08/05/2024	PERIODO UN(	ASPI6208	23				
08/05/2024	PERIODO UN(	CARCHA6261	23				
08/05/2024	PERIODO UN(	CAROLI6328	13	6	2	1	
08/05/2024	PERIODO UN(	CHUJI6317	22	1			
08/05/2024	PERIODO UN(	ESAI6309	23				
08/05/2024	PERIODO UN(	GAFA6328	23				
08/05/2024	PERIODO UN(	JUDA6338	21	2			
08/05/2024	PERIODO UN(	MALE6309	12	11			
08/05/2024	PERIODO UN(	MALUZ7759	19	4			
08/05/2024	PERIODO UN(	MOARI6309	23				
08/05/2024	PERIODO UN(	PEWA7752	19	2			
08/05/2024	PERIODO UN(	PILU8160	11	12			
08/05/2024	PERIODO UN(	QUIRI6311	5	13			
08/05/2024	PERIODO UN(	RAAD6370	20	3			
08/05/2024	PERIODO UN(	RUSE6296	23				
08/05/2024	PERIODO UN(	SALNA6309	17	6			
08/05/2024	PERIODO UN(	SOLU6311	22	1			
08/05/2024	PERIODO UN(	TOCE6328	22	1			
08/05/2024	PERIODO UN(	ZARHU7802	8	8			
08/05/2024	PERIODO UN(	FRAKA6230	22				
08/05/2024	PERIODO UN(	GODAN7479	19	3			
08/05/2024	PERIODO UN(	HUANN6249	22				
08/05/2024	PERIODO UN(	HUAJO6196	21	1			
08/05/2024	PERIODO UN(	LANE6273	14	8			
08/05/2024	PERIODO UN(	RAYO7564	19	3			
08/05/2024	PERIODO UN(	ROED7530	18	3			
08/05/2024	PERIODO UN(	RUDY7529	13	9			
08/05/2024	PERIODO UN(	SALU6224	22				
08/05/2024	PERIODO UN(	TOGRE6249	20	2			
08/05/2024	PERIODO UN(	ACAL6238	12	10			
08/05/2024	PERIODO UN(	ASZA6199	19	3	3		
08/05/2024	PERIODO UN(	BALVI6038	22				
08/05/2024	PERIODO UN(	CARES6196	12	7	2	1	
08/05/2024	PERIODO UN(	DIJA6235	20				
08/05/2024	PERIODO UN(	FELI6166	17	4	1		

Figura 37: Registro de asistencia en Excel

b) Verificar la calidad de los datos:

En esta fase se hace un análisis de la calidad de los datos, es decir analizamos si los datos tienen inconsistencias teniendo en cuenta los siguientes detalles:

Existen datos incompletos, algunos registros están vacíos o en algunos casos los datos ingresados no coinciden.

ETAPA 3: PREPARACIÓN DE LOS DATOS

En esta tercera fase se utilizó el lenguaje Python la cual nos permitió aplicar técnicas de minería de datos, realizando una visualización resumida de la estructura de la información e identificando los valores nulos, así como la información en detalle del conjunto de datos como número, etiquetas de las columnas, tipos de datos número de celdas en cada columna.

Luego de realizar la limpieza y transformación de los datos, se tuvo el registro de 2110 que corresponde a los periodos del año 2021 hasta el 2024, en dichos periodos del 2021 y 2022 se llevaron a cabo de manera virtual.

ETAPA 4: ETAPA DEL MODELADO

Los datos que se utilizan son datos de los periodos del año 2021 hasta el 2024, lo que procede a la utilización de los algoritmos de aprendizaje automático, com ose visualiza en la Figura 38.



```
# Dividir los datos en conjuntos de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X_imputed, y, test_size=0.8, random_state=42)

# Lista de modelos a comparar
models = {
    'Linear Regression': LinearRegression(),
    'Ridge Regression': Ridge(alpha=1.0),
    'Lasso Regression': Lasso(alpha=0.1),
    'Decision Tree': DecisionTreeRegressor(random_state=42),
    'Random Forest': RandomForestRegressor(n_estimators=150, random_state=42),
    'Gradient Boosting': GradientBoostingRegressor(n_estimators=150, random_state=42),
    'SVR': SVR(kernel='linear'),
    'KNN': KNeighborsRegressor(n_neighbors=5),
    'Neural Network': MLPRegressor(hidden_layer_sizes=(100,), max_iter=1000, random_state=42)
}
```

Figura 38: Lista de algoritmos a comparar en Colab

ETAPA 5: EVALUACIÓN

Una vez que se realizó la evaluación de las técnicas de Aprendizaje Automático de la base de datos, con las variables escogida en la tabla 2, para predecir el rendimiento académico, utilizando los algoritmos de Bosque Aleatorio, Gradient Boosting, Árbol de Decisión, Redes Neuronales y Maquinas de Decisiones Vectoriales, del conjunto de datos brindado, los cuales fueron analizadas mediante el lenguaje Python.

Observaremos que se mostrará que el algoritmo más optimo es el algoritmo SVM (para tipo de aprendizaje), mientras que Gradient Boosting (para la predicción de notas).

ETAPA 6: DESPLIEGUE

Los resultados obtenidos serán entregados a las autoridades de la Institución Educativa para que tomen medidas correctivas para mejorar el rendimiento académico de los estudiantes del nivel secundario.

Anexo 8: Artículo Científico

Universidad Nacional de Cañete. CHUMPITAZ Jeanice, HINOJOSA Anel

SISTEMA DE PREDICCIÓN PARA EL RENDIMIENTO ACADÉMICO EN LOS ESTUDIANTES DE LA IE MIS PRIMEROS PASOS, CAÑETE.

CHUMPITAZ Jeanice, HINOJOSA Anel

1878549112@undc.edu.pe, 1876302907@undc.edu.pe

Taller de Tesis, Escuela profesional de Ingeniería de Sistemas, Universidad Nacional de Cañete

Abstract— The objective of this research is to predict the academic performance of students at the Mis Primeros Pasos Educational Institution, Cañete. This need arose as a result of the academic performance of the students, using information from existing historical data at the Mis Primeros Pasos Educational Institution. To carry out this process, the CRISP-DM methodology was used, which consists of six stages: (1) Business compression, (2) Data compression, (3) Data preparation, (4) Modeling, (5) Evaluation, and (6) Deployment. The research approach was quantitative, applied, with a non-experimental correlational-causal design, using a sample size of 2,210 records of notes obtained through non-probabilistic convenience sampling; Document analysis was used as a technique and instrument to record historical data. As a result, it was determined that the Gradient Boosting algorithm was optimal for making grade predictions, evaluated through the mean square error, while the SVM algorithm showed greater reliability in the predictions of student learning. In conclusion, data mining techniques allowed us to significantly predict the academic performance of the students of the Mis Primeros Pasos Educational Institution, Cañete. It is recommended to implement a continuous monitoring system using advanced data mining techniques to continue evaluating and improving predictions of students' academic performance.

Keyword— Gradient Boosting, CRISP-DM methodology, SVM, mean square error

Resumen— La presente investigación tiene como objetivo predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete. Esta necesidad surgió a raíz del desempeño académico de los estudiantes, utilizando información de datos históricos existentes en la Institución Educativa Mis Primeros Pasos. Para llevar a cabo este proceso, se empleó la metodología CRISP-DM, que consta de seis etapas: (1) Comprensión del negocio, (2) Comprensión de los datos, (3) Preparación de los datos, (4) Modelado, (5) Evaluación, y (6) Despliegue. El enfoque de la investigación fue cuantitativo, de tipo aplicado, con un diseño no experimental de tipo correlacional-causal, utilizando un tamaño de muestra de 2210 registros de notas obtenidos mediante muestreo no probabilístico por conveniencia; se utilizó el análisis documental como técnica e instrumento para registrar los datos históricos. Como resultado, se determinó que el algoritmo Gradient Boosting fue óptimo para realizar predicciones de notas, evaluado a través del error cuadrático medio, mientras que el algoritmo SVM mostró una mayor fiabilidad en las predicciones del aprendizaje de los estudiantes. En conclusión, las técnicas de minería de datos permitieron prever significativamente el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete. Se recomienda implementar un sistema de monitoreo continúo utilizando técnicas avanzadas de minería de datos para seguir evaluando y mejorando las predicciones del rendimiento académico de los estudiantes.

Palabra clave— Gradient Boosting, metodología CRISP-DM, SVM, error cuadrático medio

I. INTRODUCCIÓN

En el mundo actual, la digitalización ha provocado un enorme aumento en la cantidad de datos disponibles, pero este crecimiento no siempre conduce a una mayor generación de conocimiento, como lo argumentan. En numerosos sectores académicos, científicos y empresariales, el procesamiento de datos utilizando métodos tradicionales frecuentemente resulta ser difícil, extremadamente laborioso y produce resultados superficiales e insatisfactorios [1]. Uno de los temas discutidos en el ámbito educativo es el rendimiento académico, que se enfoca principalmente en los logros o dificultades que enfrentan los estudiantes. Dado que los alumnos son pilares fundamentales de las instituciones educativas y de la sociedad en general, surge la necesidad de prever aquellos estudiantes que pueden tener un periodo educativo exitoso o complicado. Esto permitiría a la universidad ofrecer alternativas para mejorar y respaldar el progreso académico de los estudiantes [2].

Es por ello que el avance significativo de la analítica ha penetrado en el ámbito educativo, donde ahora se maneja un extenso conjunto de datos que incluye información sobre los participantes del proceso educativo. Es en este punto donde la ingeniería puede desempeñar un rol crucial al contribuir a resolver varios aspectos académico-administrativos, como la mejora del aprendizaje, la

reducción de la deserción, la disminución del abandono y la optimización del rendimiento académico [3].

En este contexto, la aplicación de tecnologías enfocadas en la extracción de datos emerge como una solución eficiente para el análisis de información, y la imperiosa tarea de elevar el estándar educativo plantea diversas estrategias para cambiar de forma dinámica la concepción de la enseñanza [4].

A nivel global, en algunas instituciones educativas de Ecuador se llevan a cabo sesiones de evaluación al finalizar un examen parcial, trimestre o año escolar para analizar el desempeño académico de los estudiantes. Este análisis se centra principalmente en determinar cuántos estudiantes han tenido éxito y cuántos no han alcanzado los estándares requeridos en cada asignatura. Sin embargo, este enfoque se limita a un análisis descriptivo básico utilizando técnicas estadísticas convencionales. Esto implica que se recopila información sobre la cantidad de estudiantes con bajo rendimiento, pero no se investigan en profundidad las causas fundamentales que provocan estos resultados negativos [5].

A nivel nacional, en Perú se ha observado que el 33.1% de los estudiantes obtienen un promedio de nota igual o menor a 10.5 en modalidad presencial, mientras que este porcentaje aumenta significativamente al 55.7% bajo la modalidad virtual. Estas estadísticas son motivo de preocupación, ya que indican que casi la mitad de los alumnos matriculados muestran un desempeño académico deficiente. [6].

Tal es la situación de la Institución Educativa Mis Primeros Pasos, ubicada en el distrito de Mala de la Provincia de Cañete brinda servicios educativos, lo cual consta de tres niveles: inicial, primaria y secundaria.

La institución educativa, que apuesta por brindar servicios educativos, ha enfrentado algunos desafíos en materia de calidad de la educación, pues el rendimiento académico de los estudiantes el año pasado fue insuficiente en comparación con el momento de la pandemia de coronavirus de 2019.

Dicha Institución carece de una información detallada, pero sí dispersa manualmente ya sea: archivos digitales, seguimiento escolar en los docentes de la Institución. Este problema en los estudiantes podría afectar en su conducta o entre otros factores y sobre todo la pérdida del año escolar.

Sin embargo, el propósito de este estudio es predecir el desempeño académico de los estudiantes utilizando información de datos históricos existentes de los estudiantes en la institución educativa Mi primer paso.

A. Formulación del problema

a) Problema general

¿Cómo predice un sistema el rendimiento académico de los estudiantes de la institución educativa Mis Primeros Pasos, Cañete?

b) Problemas específicos

- ¿Qué indicadores de impacto pueden ayudar a determinar los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete?
- ¿Cuál es el algoritmo óptimo para predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete?
- ¿Como la predicción de rendimiento académico ayuda a detectar a los estudiantes que necesitan aprendizaje personalizado en la Institución Educativa Mis Primeros Pasos, Cañete?

B. Objetivos

a) Objetivo general

Determinar la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

b) Objetivos específicos

- Determinar los indicadores de impacto que ayuda a reconocer los factores que influyen en el rendimiento académico de los estudiantes en la institución educativa Mis Primeros Pasos, Cañete.
- Determinar el algoritmo óptimo que genere un modelo para predecir el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.
- Determinar una predicción al rendimiento académico que ayudará con el aprendizaje personalizado a los estudiantes de la Institución Educativa Mis Primeros Pasos.

C. Marco conceptual

Sistema de Predicción. Un sistema de predicción en el contexto del rendimiento académico es una aplicación del aprendizaje automático que utiliza datos históricos y variables relevantes para prever cómo un estudiante o grupo de estudiantes pueden desempeñarse en el futuro. Este tipo de sistema utiliza algoritmos y modelos matemáticos para analizar patrones en los datos y generar predicciones basadas en esos patrones. Los datos utilizados pueden incluir información como el historial académico previo, características demográficas, resultados de pruebas estandarizadas, asistencia, entre otros [17].

Rendimiento Académico. El rendimiento académico es aquel factor supuesto que da a conocer la capacidad que tiene un estudiante, todo ello es evaluado mediante pruebas, actividades y otros que se van presentando en el transcurso del año escolar. Por otro lado, en algunos casos, se puede percibir que en los adolescentes es un cambio al cual deben adaptarse dado que el estudio es más fuerte y puede ser más estresante para algunos.[21]

D. Dimensiones

Técnicas predictivas: Es avanzada debido a su análisis sofisticado, que emplea datos históricos para el aprendizaje y datos nuevos para las predicciones. Dentro de su estructura de análisis se utilizan técnicas estadísticas y algoritmos de aprendizaje, los cuales son empleados para identificar regularidades, configuraciones o tendencias, con la finalidad de ofrecer una estimación confiable sobre un acontecimiento venidero [13].

En el presente proyecto de investigación para la variable dependiente se consideraron como dimensiones:

- Nivel Económico: En este estudio, se exploró la dimensión del nivel económico como medida económica y social de una familia. Se utilizaron como indicadores el tipo de colegio y la escala de pensión, considerando que el estudiante no recibe ningún tipo de apoyo externo [21].
- Nivel demográfico: En la presente tesis, se examinó la dimensión de nivel demográfico ya que es necesario obtener informaciones de los atributos del estudiante como la edad, sexo, lugar de residencia, entre otros [14].
- Nivel Académico: En este trabajo de investigación, se evaluó la dimensión relacionada con las habilidades del estudiante que se han desarrollado y mejorado durante el proceso de enseñanza-aprendizaje. Este proceso permite al estudiante alcanzar un nivel de rendimiento y logros académicos durante un periodo o semestre determinado [16].

II. TRABAJOS RELACIONADOS

Según Quijije y Maldonado en el año 2023, en su artículo científico titulado "TÉCNICA DE MINERIA DE DATOS PARA PROCESOS EDUCATIVOS EN ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES BASADO EN UN MODELO PREDICTIVO", el propósito fue desarrollar un modelo predictivo utilizando técnicas de exploración de datos aplicadas a procesos educativos de estudiantes con necesidades educativas especiales en la Unidad Educativa "Plan Internacional". En este estudio se utilizó la metodología CRISP-DM. Como resultado, se resaltó el algoritmo J48 como sólido para la construcción de modelos predictivos del desempeño académico, con potencial aplicabilidad en otros entornos educativos. En resumen, el uso de técnicas de exploración de datos en el ámbito educativo sigue evolucionando y adaptándose a los requisitos específicos. Es fundamental obtener datos que validen los procesos y la retención del conocimiento en instituciones educativas, lo que subraya la necesidad continua de mejorar los análisis y enfoques de la exploración de datos, especialmente en aplicaciones prácticas y tareas de predicción y descripción que respalden la toma de decisiones educativas. [4].

Según Aronés en el año 2021, llevó a cabo un estudio titulado "PREDICCIÓN DEL RENDIMIENTO ACADÉMICO BASADO EN MACHINE LEARNING, ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS, AYACUCHO 2021". El objetivo fue

diseñar un modelo de aprendizaje automático (Machine Learning) con el propósito dar predicción en el rendimiento académico de los universitarios de la Escuela Profesional de Ingeniería de Sistemas. Optó por una investigación de tipo transversal, observacional y retrospectivo con nivel descriptivo y diseño no experimental. La población se compuso de los registros de los estudiantes de dicha Escuela Profesional de los semestres 2016-II hasta 2019-I, lo cual se tomaron los mismos datos como muestra. El resultado obtenido fue favorable aplicando el algoritmo de regresión logística que ayudó a la predicción con un 73.6% de acuerdo a la métrica de validación de la curva ROC. Se concluye que el trabajo de investigación brinda una mejor probabilidad predictiva y toma de decisiones adecuadas en el rendimiento académico de los estudiantes de la Escuela Profesional [13].

Por otro lado, a nivel internacional, Gonzabay en el año 2023 realizó un estudio titulado "Evaluación del desempeño académico de la U.E John F. Kennedy ubicada en la parroquia José Luis Tamayo del cantón Salinas mediante técnicas de minería de datos", con el objetivo principal de aplicar la minería de datos para identificar patrones en el rendimiento académico de los estudiantes y determinar así qué estrategias de aprendizaje deben mejorarse. A lo cual tuvo como resultado predictivo gracias al dashboard interactivo haciendo uso del Power BI, en el cual las partes interesadas hicieron el monitoreo respectivo de la información identificando los factores que influyen en el rendimiento académico llegando a la búsqueda de nuevas estrategias de aprendizaje [15].

III. METODOLOGÍA

a. Tipo de investigación

El estudio aplicado [7] es de tipo de investigación científico, está basado en la identificación y análisis de los procesos prácticos, formulación de la solución y el diseño del plan piloto para brindar soluciones en diversos campos y sectores para la toma de decisiones y una mejor eficacia brindando soluciones efectivas.

Es aplicada en base a que se determinará la predicción del rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.

b. Nivel de investigación

El nivel explicativo [8] permite identificar factores de causa y efecto de las relaciones entre variables, para el desarrollo de teorías y modelos que explican el fenómeno estudiado. En este caso, el nivel de investigación es explicativo, puesto que la necesidad se basará en identificar los cambios del sistema predictivo que influyen en la toma de decisiones del rendimiento académico de los estudiantes

c. Método de investigación

Teórico-analítico-sintético: se busca caracterizar la relación que existe entre los indicadores económicos, demográficos y rendimiento de los estudiantes del nivel secundaria de la Institución Educativa Mis Primeros Pasos.

Teórico lógico-histórico: evaluación de los estudiantes del nivel secundario de la Institución Educativa Mis Primeros Pasos mediante datos históricos influyentes en su rendimiento académico como: datos económicos, demográficos y rendimiento.

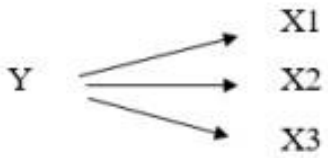
Teóricos por modelación: se da una descomposición operacional ya que se busca saber la relación que existe en los indicadores económicos, demográficos y de rendimiento de los estudiantes del nivel secundario de la Institución Educativa Mis Primeros Pasos

d. Diseño de investigación

El diseño No-Experimental (transeccional) del tipo correlacional causal [7] no implica la manipulación intencional de variables independientes por parte del investigador. En su lugar, se observan y registran las variables según se presentan naturalmente en los participantes o en las situaciones bajo estudio. Los estudios transversales recopilan datos en un

solo punto en el tiempo, proporcionando una imagen instantánea de las relaciones entre variables en ese momento.

Donde se busca la relación que existe en el rendimiento académico con los indicadores económicos, demográficos y de rendimiento de cada uno de los estudiantes de la Institución.



Donde:
Y: Rendimiento académico
X1: Indicador económico
X2: Indicador demográfico
X3: Indicador de rendimiento

IV. RESULTADOS

Se presentan los resultados que se llegaron a obtener en el proceso de investigación, donde se desarrolló el modelo predictivo más óptimo para la predicción siendo este el objetivo principal.

RESULTADOS DE LA PRUEBA DE MODELOS DE PREDICCIÓN

La Figura 10 muestra mediante un gráfico de barras la comparación del error cuadrático medio de los diferentes modelos, donde:

- Existen 4 modelos con un error cuadrático medio cercano a 4, siendo estos: Linear Regression, Ridge Regression, Lasso Regression y Neural Netwokrd.
- Hay 3 modelos con un error cuadrático medio cercano a 4.5, siendo estos: Random Forest, Support Vector Machine y K – Neighbors Regressor
- Un modelo con error cuadrático medio cercano a 5.5, siendo este Decision Tree.
- Un modelo con un error cuadrático medio alrededor de 3.5, siendo este Gradient Boosting.
- Teniendo como modelo predictivo el Gradient Boosting, el cual tiene una mejor precisión en el rendimiento académico en comparación a los otros algoritmos.

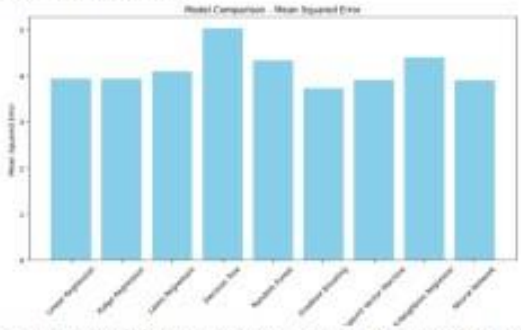


Figura 1: Model Comparison - Mean Squared Error

Para la elección del algoritmo óptimo se divide los datos en conjuntos de entrenamiento y así poder hacer la prueba de cada uno de estos; donde se inicia el entrenamiento y evaluación de cada modelo como se puede apreciar en la Figura 13:

```
Training Logistic Regression...
Logistic Regression Accuracy: 0.9447077409162717
              precision    recall  f1-score   support

           1       0.95      1.00      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.94       633
   macro avg       0.47      0.50      0.49       633
weighted avg       0.90      0.94      0.92       633

--------------------------------------
Training Decision Tree...
Decision Tree Accuracy: 0.9368088467614534
              precision    recall  f1-score   support

           1       0.95      0.99      0.97       599
           2       0.00      0.00      0.00        34

    accuracy                           0.94       633
   macro avg       0.47      0.49      0.48       633
weighted avg       0.89      0.94      0.92       633

--------------------------------------
Training Random Forest...
Random Forest Accuracy: 0.9447077409162717
              precision    recall  f1-score   support

           1       0.95      0.99      0.97       599
           2       0.43      0.09      0.15        34

    accuracy                           0.94       633
   macro avg       0.69      0.54      0.56       633
weighted avg       0.92      0.94      0.93       633
```

Figura 2: Prueba de algoritmo óptimo

Como resultado de los algoritmos se puede observar los detalles de cada uno luego de ser evaluados, donde se percibe que la métrica que arroja es muy cercana a 1 siendo este el algoritmo SVM con una exactitud más confiable que los otros algoritmos. Por ello, se

considera un algoritmo óptimo para trabajar las predicciones.

Prueba de hipótesis con los indicadores de nota al final de trimestre y asistencia.

En el mapa de calor se pudo obtener el impacto que puede tener cada dato obtenido, dando a conocer cuáles son los que ayudarán a predecir el rendimiento académico del estudiante. Se puede observar en la Figura 20 que la Nota anterior es la que tendrá más uso para realizar una predicción real.

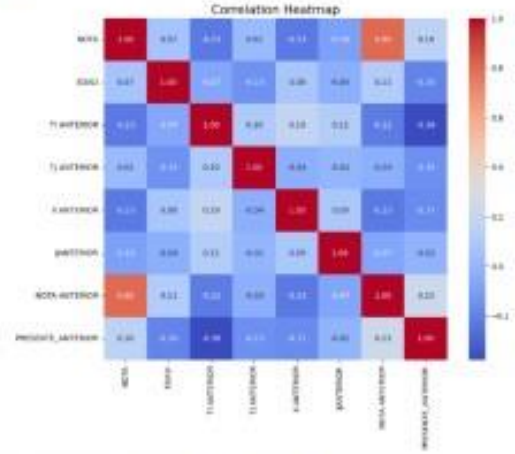


Figura 3: Mapa de calor de correlación

V. CONCLUSIONES

1. Se determinó que las técnicas de minería de datos predicen significativamente el rendimiento académico de los estudiantes de la Institución Educativa Mis Primeros Pasos, Cañete.
2. Se identificó mediante un mapa de calor el indicador clave para predecir el rendimiento académico, donde la nota es la que más influencia tiene.
3. Se determinó que el algoritmo con mejor precisión para predecir el rendimiento académico de los estudiantes es el modelo Gradient Boosting que presenta un rendimiento general y mejor para realizar la predicción que se quiere, teniendo un MSE de 3.724235 y un R^2 de 0.553136, lo que indica que es un modelo que se ajusta a los datos en la investigación en comparación a los otros algoritmos.
4. Se determinó que el algoritmo SVM tuvo una mejor confiabilidad en las predicciones y con mayor exactitud que los otros algoritmos, lo cual se pudo identificar a los estudiantes que necesita un aprendizaje personalizado de la materia.

VI. REFERENCIAS

[1] D.C. Chamorro-Sangoquiza, A.M. Vargas-Muñoz, A.C. Umaquinga-Criollo, M.A. Becerra y D.H. Peluffo-Ordoñez, "Estudio comparativo de técnicas de mineria de datos para develar patrones de desempeño académico en enseñanza media", *RISTI*, pp. 455-468, junio de 2020.

[2] J.D. Garcia Dionisio, "Machine Learning para predecir el rendimiento académico de los estudiantes universitarios", tesis de Licenciado, Universidad César Vallejo, Lima, UCV, Perú, 2021.

[3] L.E. Contreras, H.J. Fuentes y J.I. Rodriguez, "Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático", *SCIELO*, vol. 13(5), pp. 233-246, octubre de 2020.

[4] H.B. Quijije Quiroz y K. Maldonado Zuñiga, " TÉCNICA DE MINERÍA DE DATOS PARA PROCESOS EDUCATIVOS EN ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES BASADO EN UN MODELO PREDICTIVO", *PENTACIENCIAS*, vol. 5, pp. 205-217, septiembre de 2023.

[5] M.A. Gonzabay Panchana, "Evaluación del rendimiento académico de la U.E John F. Kennedy situada en la parroquia José Luis Tamayo del cantón Salinas aplicando técnicas de mineria de datos", tesis de Maestria, Universidad Estatal Peninsula de Santa Elena, Santa Elena, UPSE, Ecuador, 2023.

[6] L.E. Contreras, H.J. Fuentes y J.I. Rodriguez, "Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingenieria, mediante aprendizaje automático", *SCIELO*, vol. 13(5), pp. 233-246, octubre de 2020.

[7] R. Hernández Sampieri, C. Fernández Collado y P. Baptista Lucio. *Metodología de la Investigación*. México: McGrawHill, 2019.

[8] J. Arias-Gonzales, *DISEÑO Y METODOLOGÍA DE LA INVESTIGACIÓN*, Arequipa: ResearchGate, 2021.

[9] J.S. Salto-Mero y M.R Cruz-Felipe, "Análisis del Rendimiento Académico de Estudiantes de las Carreras Economía y Turismo con Power BI en los Periodos (2021)", *Digital Publisher*, vol. 5, pp. 762-772, enero de 2024.

[10] L. Quiñones Huatangari y Y.L. Carrasco Vega, " Rendimiento académico empleando mineria de datos", *Espacios*, vol. 41(44), pp. 277-285, noviembre de 2020.

[11] R. Rojas Pari, "Modelo de Aprendizaje Automático Supervisado para Identificar Patrones de Bajo Rendimiento Académico en los Ingresantes al Instituto de Educación Superior Pedagógico Público – Juliaca ". tesis de Licenciado, Universidad Peruana Unión, Juliaca UPEU, Perú, 2021.

[12] W.G. Paucar Palomino, "MODELO DE APRENDIZAJE AUTOMATICO PARA LA PREDICCION DEL RENDIMIENTO ACADEMICO DE LOS ALUMNOS INGRESANTES EN LA FACULTAD DE INGENIERIA EN INDUSTRIAS ALIMENTARIAS DE LA UNAS", tesis de Licenciado, Universidad Nacional Agraria de la Selva, Tingo Maria, UNAS, Perú, 2023.

[13] E. Aronés Ayala, " PREDICCIÓN DEL RENDIMIENTO ACADÉMICO BASADO EN MACHINE LEARNING, ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS, AYACUCHO 2021", tesis de Licenciado, Universidad Nacional de San Cristóbal de Huamanga, Ayacucho, UNSCH, Perú, 2021.

[14] N.H. Luna Reyes, "IMPLEMENTACIÓN DE UN SISTEMA DE PREDICCIÓN DEL RENDIMIENTO ACADÉMICO DE LOS ESTUDIANTES DE INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD NACIONAL JOSÉ MARÍA ARGUEDAS UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS PARA LA ADECUADA TOMA DE DECISIONES", tesis de Licenciado, Universidad Nacional José María Arguedas, Apurimac, UNAJMA, Perú, 2020.

[15] M.A. Gonzabay Panchana, "Evaluación del rendimiento académico de la U.E John F. Kennedy situada en la parroquia José Luis Tamayo del cantón Salinas aplicando técnicas de mineria de datos.", tesis de Magister, Universidad Estatal Península de Santa Elena, Santa Elena, UPSE, Ecuador, 2023.

[16] L.A. Mantilla Galarza, " Análisis del rendimiento académico de los estudiantes de la Unidad Educativa "Aristóteles School" aplicando herramientas de Business Intelligence y mineria de datos para mejorar el proceso de enseñanza aprendizaje.", tesis de Magister, Universidad de las Fuerzas Armadas, Santa Elena, ESPE, Ecuador, 2023.

[17] «Qué es Modelo predictivo,» ARIMETRICS, [En línea]. Available: https://www.arimetrics.com/glosario-digital/modelo-predictivo. [Último acceso: 23 mayo 2024].

[18] K. J. Paredes-Robles, J. Vélez-Pincay y L. R. Marin-Llaver, «Estrategia Educativa para mejorar el rendimiento académico en estudiantes de bachillerato.,» 15 septiembre 2023. [En línea]. [Último acceso: mayo 2024].

[19] prom peru, SOFTWARE Mineria de Datos y Almacenamiento, 2021.

[20] S. PIO ALVAREZ, «MINERÍA DE CALIDAD DE DATOS: APLICACIÓN DE TÉCNICAS DE MINERÍA DE DATOS PARA LA EVALUACIÓN DE LA CALIDAD DE LOS DATOS,» 2018.

[21] A. J. VILLAO BALÓN, «Aplicación de técnicas de minería de datos para predecir el desempeño académico de los estudiantes de la escuela 'Lic. Angélica Villón L.,» 2021

[22] A. Müller. y S. Guido (2019). Introduction to Machine Learning with Python. Boston: OREILLY.

[23] M. Pojon (2020). Using Machine Learning to predict student performance. University of Tampere

[24] M. Orallo, C. Ramírez (s. f.) "Minería de Datos y Extracción de Conocimiento de Bases de Datos".

[25] A. Rico Páez (24 de Enero-junio de 2022). "Modelos preditivos progressivos de desempenho acadêmico de estudantes". 12(24).

[26] E. Taya Acosta. "Modelo de mineria de datos para evaluar el efecto del uso del aula virtual sobre el rendimiento academico de los estudiantes de la facultad de ingenieria de la universidad nacional Jorge Basadre Grohmann de Tacna, en Tiempos de pandemia,2020". tesis de Magister, Universidad Nacional Jorge Basadre Grohmann, UNJBG, Tacna, 2021.

[27] J. Puga de la Cruz, & R. Torres Monzon, "REDES NEURONALES ARTIFICIALES PARA PRONOSTICAR EL RENDIMIENTO ACADÉMICO DE ALUMNOS DE INGENIERÍA DE SISTEMAS E INFORMÁTICA DE LA UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA". Repositorio Institucional Digital UNAP, Amazonas, 2023

[28] A. Oviedo Carrascal, & J. Jiménez Giraldo, "Minería de datos educativos: Análisis del desempeño de estudiantes de ingeniería en las pruebas saber-pro. Dialnet, 1-13. 2019

[29] A. Páez Rico, & N. Ramírez Gaytán, "Predicción del rendimiento académico utilizando las primeras actividades académicas de estudiantes universitarios y técnicas de aprendizaje automático". (31 de agosto de 2022)

Anexo 9: Constancia de similitud

UNIVERSIDAD NACIONAL DE CAÑETE
Ley de Creación N° 29488
FACULTAD DE INGENIERÍA
UNIDAD DE INVESTIGACION DE LA FACULTAD DE INGENIERIA
"Año del Bicentenario, de la Consolidación de nuestra Independencia
Y de la conmemoración de las heroicas batallas de Junín y Ayacucho"

San Vicente de Cañete, 20 de julio del 2024

CONSTANCIA DE SIMILITUD DE INFORME FINAL DE TESIS N° 009-2024

Titulo del Trabajo de Investigación: "Sistema de pronóstico para el rendimiento académico y conductual en la IE Mis Primeros Pasos, Cañete"

Autor(es): Chumpitaz Luna Jeanice Daniela
Hinojosa Rojas Anel Cecilia


JEANICE DANIELA CHUMPITAZ LUNA - ANEL CECILIA HINOJOSA ROJAS

INFORME DE ORIGINALIDAD

8%	8%	4%	3%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	repositorio.undc.edu.pe Fuente de Internet	4%
2	repositorio.unas.edu.pe Fuente de Internet	1%
3	repositorio.unajma.edu.pe Fuente de Internet	1%
4	593dp.com Fuente de Internet	1%

Excluir citas Activo
Excluir bibliografía Activo
Excluir coincidencias < 1%

Por lo tanto, se otorga esta constancia a solicitud del interesado(a) para los fines que estime conveniente.

Mg. Amanda Durán Carhuamaca
(e) Directora de la Unidad de Investigación
Facultad de Ingeniería – UNDC

FIRMA DIGITAL
Firmado digitalmente por: DURAN CARHUAMACA AMANDA FIR 20114878 hard
Motivo: En señal de conformidad
Fecha: 20/07/2024 13:52:08-0500

Anexo 10: Autorización de inclusión en el repositorio de la biblioteca

AUTORIZACIÓN DE INCLUSIÓN EN EL REPOSITORIO DE LA BIBLIOTECA

Yo, **HINOJOSA ROJAS ANEL CECILIA**, titular del documento de identidad número **76302907**, estudiante de la carrera **INGENIERÍA DE SISTEMAS** de la Universidad Nacional de Cañete (UNDC), por medio de la presente, en calidad de autor y actuando en representación de mis co-autores, autorizo a la Biblioteca Central de la UNDC a incluir y difundir la tesis titulada **"Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete"**, que he presentado para obtener el grado de **Ingeniero de sistemas**, en su repositorio digital y otros medios electrónicos que la universidad considere pertinentes.

A través de esta autorización:

1. Reconozco que los derechos de autor de la tesis me pertenecen y que cualquier uso indebido por parte de terceros es responsabilidad exclusiva de quienes lo realicen.
2. Permito que la tesis pueda ser consultada por usuarios internos y externos de la Universidad y que se realicen copias de seguridad, migraciones de formato y otras acciones necesarias para su preservación y acceso.
3. Acepto que esta autorización no implica la renuncia a ningún derecho que me corresponda como autor de la tesis, incluidos los derechos de reproducción, distribución y comunicación pública.
4. Entiendo que la Universidad se compromete a utilizar la tesis exclusivamente con fines académicos y no comerciales.
5. Declaro que la información proporcionada en este documento es veraz y completa. Me hago responsable por cualquier inexactitud o falsedad en la misma.

Fecha: 11 de Julio de 2024

Atentamente,

Firma:



Av. Mariscal Benavides 1370, Sede Académica: Casa de la Cultura.
Código Postal 15701. San Vicente, Cañete, Lima, Perú | www.undc.edu.pe

Anexo 11: Autorización de repositorio institucional de la UNDC

UNIVERSIDAD NACIONAL DE CAÑETE	POLÍTICAS Y REGLAMENTO DEL REPOSITORIO INSTITUCIONAL DEACCESO ABIERTO		
	Código: R-M02-VRI-004	Revisión: 01	Fecha: Marzo-2023

AUTORIZACIÓN DE PUBLICACIÓN EN EL REPOSITORIO INSTITUCIONAL DE LA UNDC

I. INFORMACIÓN DEL DOCUMENTO

1.1 Título del documento:

Sistema de predicción para el rendimiento académico en los estudiantes de la IE Mis Primeros Pasos, Cañete.

1.2 Descripción (Seleccionar)

Tesis de Pregrado	X	Libro	
Tesis de Maestría		Capítulo de libro	
Tesis de Doctorado		Monografía	
Trabajo de Investigación		Artículo	
Tesis de Segunda Especialidad		Conferencia	
Trabajo de Suficiencia Profesional		Programa informativo	
Trabajo académico		Datos	

1.3 Autores del documento

Apellidos	Nombres	Correo	DNI	ORCID
Chumpiaz Luna	Jeanice Daniela	1878549112@undc.edu.pe	78549112	orcid.org/0009-0007-8447-624X
Hinojosa Rojas	Anel Cecilia	1876302907@undc.edu.pe	76302907	orcid.org/0000-0002-3027-1405

1.4 Asesor (es)

Apellidos	Nombres	Correo	DNI	ORCID
Sanchez Castillo	Eddye Arturo	e_sanchez@undc.edu.pe	06011207	orcid.org/0000-0002-9534-6156
Almidón Ortiz	Carlos Alcides	calmidon@undc.edu.pe	20066294	orcid.org/0000-0003-1055-9724

1.5 Centro de Investigación

Facultad	De Ingeniería
Escuela Profesional	De Ingeniería de Sistemas

II. ORIGINALIDAD DEL TRABAJO PRESENTADO

Con la presentación de esta ficha, el(los) autor(es) señalan expresamente que la obra es original, ya que sus contenidos son producto de su directa contribución intelectual. Se reconoce también que todos los datos y las referencias a materiales ya publicados están debidamente identificados con su respectivo crédito e incluidos en las notas bibliográficas y en las citas que se destacan como tal.

III. AUTORIZACIÓN DEL TRABAJO PRESENTADO

Por medio del presente documento, afirmo y garantizo(amos) ser el (los) legítimo (s), titular(es) de todos los derechos de propiedad intelectual sobre los entregables, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los "Contenidos" que serán incluidos en el Repositorio Institucional de la Universidad Nacional de Cañete (en adelante, la "Universidad").

Autorizo a la Universidad a publicar los Contenidos en el Repositorio Institucional de la Universidad y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con su Repositorio Institucional. Autorizo expresamente toda consulta y uso de los contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.

En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determinen, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

IV: PUBLICACIÓN DIFERIDA

Este ítem solo es de interés para los autores que han decidido la publicación diferida de sus documentos. Caso contrario, obvia este apartado.

Autorización	Motivo (marcar)		Firma
Solicito la publicación diferida del documento depositado en el repositorio, por 12 meses. (indicar el tiempo que solicita)	Exclusividad de revista, editor		
	Por patente		
	Secreto o seguridad nacional		
	Por otras razones		

Cualquier motivo utilizado, el solicitante debe acreditar con documento su pedido (subir junto con la autorización la evidencia que sustenta su solicitud).

Nosotras, Chumpitaz Luna Jeanice Daniela con DNI N° 78549112 e Hinojosa Rojas Anel Cecilia con DNI N° 76302907 en calidad de autores y actuando en representación de mis co-autores, autorizo la publicación del documento indicado en el punto 1, bajo las condiciones indicadas en el punto 2 y 3, dejando constancia que el archivo digital que estoy entregando a la Universidad Nacional de Cañete, como parte del proceso de obtención del título profesional o grado académico, contiene la versión final del documento sustentado y aprobado por el jurado.



11/07/2024

Firma

Fecha